GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:02:27 ; Search time 51.308 Seconds
(without alignments)
699.375 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	618	90.0	123	3	AAY95426	Aay95426 Human hig
2	618	90.0	304	6	ABP70801	Abp70801 Human ext
3	618	90.0	322	6	ABP70799	Abp70799 Human ext
4	618	90.0	329	6	ABU92044	Abu92044 Human pro
5	618	90.0	358	6	ABP70800	Abp70800 Human ext
6	618	90.0	390	6	ABU99149	Abu99149 Novel hum
7	618	90.0	398	6	ABU99143	Abu99143 Novel hum
8	618	90.0	427	8	ADE76864	Ade76864 Human pro
9	618	90.0	615	6	ABU99144	Abu99144 Novel hum
10	618	90.0	626	5	ABB78707	Abb78707 Human hig
11	618	90.0	644	4	ABG21101	Abg21101 Novel hum
12	618	90.0	644	5	ABB78710	Abb78710 Human hig
13	618	90.0	644	6	ABU99150	Abu99150 Novel hum
14	618	90.0	644	6	ABU99145	Abu99145 Novel hum
15	586	85.3	122	3	AAB37447	Aab37447 Human kin
16	585	85.2	435	4	ABG21105	Abg21105 Novel hum
17	556.5	81.0	117	2	AAR33350	Aar33350 Domaine 3
18	440	64.0	436	1	AAP40257	Aap40257 Bradykini
19	413	60.1	434	1	AAP40633	Aap40633 Bradykini
20	411	59.8	357	6	ABR41202	Abr41202 Human DIT
21	388	56.5	235	5	ABG60077	Abg60077 Human DIT
22	320.5	46.7	248	4	ABG21102	Abg21102 Novel hum
23	316	46.0	369	4	ABG21099	Abg21099 Novel hum
24	190	27.7	305	4	ABG21100	Abg21100 Novel hum
25	171.5	25.0	167	2	AAW98907	Aaw98907 Mouse IMC
26	166	24.2	32	3	AAY95418	Aay95418 Anti-angi
27	163.5	23.8	126	3	AAB37445	Aab37445 Human cys
28	163.5	23.8	145	2	AAW32323	Aaw32323 Mature hu
29	163.5	23.8	145	2	AAW31902	Aaw31902 Human cys
30	163.5	23.8	145	2	AAY25708	Aay25708 Human cys
31	163.5	23.8	145	4	AAE02410	Aae02410 Human cys
32	163.5	23.8	145	4	AAE04439	Aae04439 Human cys
33	163.5	23.8	145	7	ADD14098	Add14098 Human src
34	163.5	23.8	167	2	AAY02287	Aay02287 Secreted
35	163	23.7	167	7	ADA45154	Ada45154 Human pol
36	163	23.7	178	2	AAW69734	Aaw69734 Human cys
37	161	23.4	32	3	AAY95408	Aay95408 Anti-angi
38	157	22.9	122	3	AAB37446	Aab37446 Human kin
39	154.5	22.5	167	2	AAW98910	Aaw98910 Mouse IMC
40	153	22.3	27	3	AAY95425	Aay95425 Anti-angi
41	144.5	21.0	121	3	AAY81200	Aay81200 Human mut
42	144.5	21.0	128	3	AAY81189	Aay81189 Human mut
43	143.5	20.9	121	3	AAY81198	Aay81198 Human mut
44	143.5	20.9	128	3	AAY81187	Aay81187 Human mut
45	142.5	20.7	118	3	AAY81218	Aay81218 Bovine mu

ALIGNMENTS

RESULT 1
AAY95426
ID AAY95426 standard; peptide; 123 AA.
XX
AC AAY95426;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human high mol.wt. kininogen domain 3.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic; therapy;
KW human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US028465.
XX
PR 16-DEC-1998; 98US-0112427P.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI Mccrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain 3
PT analog.
XX
PS Disclosure; Page 4; 44pp; English.
XX
CC The present sequence is that of domain 3 of human high mol.wt. kininogen
CC (HK). The invention provides peptides (see AAY95405-24) that are
CC analogues of certain sites in the HK domain 3, specifically Asn275-
CC Lys282, Cys246-Cys249, Leu331-Tyr338 and Tyr299-Ser314. The peptides, in
CC which native Cys residues may be replaced by Ala residues, inhibit
CC endothelial cell proliferation and may also induce endothelial cell
CC apoptosis. Compositions including the peptides are used in claimed
CC methods for inhibiting angiogenesis, inhibiting endothelial cell
CC proliferation, and inducing endothelial cell apoptosis. Cancer,

CC rheumatoid arthritis, and ocular disorders characterized by undesired
CC vascularization of the retina are treated
XX
SQ Sequence 123 AA;

Query Match 90.0%; Score 618; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 5.5e-63;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
Db 1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
Db 61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 116

RESULT 2
ABP70801
ID ABP70801 standard; protein; 304 AA.
XX
AC ABP70801;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extradellular messenger, EXMES-28.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42388.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX
PS Claim 1; Page 207; 224pp; English.
XX
CC The present invention relates to novel human extracellular messenger
CC proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC for diagnosing or treating a disease or condition associated with
CC decreased expression or overexpression of functional EXMES e.g.
CC autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC cancer
XX
SQ Sequence 304 AA;

Query Match 90.0%; Score 618; DB 6; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.9e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
Db 130 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 189

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
Db 190 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 245

RESULT 3
ABP70799
ID ABP70799 standard; protein; 322 AA.
XX
AC ABP70799;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extracellular messenger, EXMES-26.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42386.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX
PS Claim 1; Page 205-206; 224pp; English.
XX
CC The present invention relates to novel human extracellular messenger
CC proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC for diagnosing or treating a disease or condition associated with
CC decreased expression or overexpression of functional EXMES e.g.
CC autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC cancer
XX
SQ Sequence 322 AA;

```
  Query Match             90.0%;  Score 618;  DB 6;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 207

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 263
```

RESULT 4
ABU92044
ID ABU92044 standard; protein; 329 AA.
XX
AC ABU92044;
XX
DT 15-JUL-2003 (first entry)
XX
DE Human protein modification and maintenance molecule-24 (PMMM-24).
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR N-PSDB; ACA92439.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 1; Page 264-265; 311pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention
XX
SQ Sequence 329 AA;

```
  Query Match             90.0%;  Score 618;  DB 6;  Length 329;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 214

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 270
```

RESULT 5
ABP70800
ID ABP70800 standard; protein; 358 AA.
XX
AC ABP70800;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extracellular messenger, EXMES-27.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42387.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX

```
PS   Claim 1; Page 206; 224pp; English.
XX
CC   The present invention relates to novel human extracellular messenger
CC   proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC   (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC   for diagnosing or treating a disease or condition associated with
CC   decreased expression or overexpression of functional EXMES e.g.
CC   autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC   cancer
XX
SQ   Sequence 358 AA;

  Query Match             90.0%;  Score 618;  DB 6;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 243

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 299


RESULT 6
ABU99149
ID   ABU99149 standard; protein; 390 AA.
XX
AC   ABU99149;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Novel human GPCR related protein NOV12g.
XX
KW   Human; G-protein coupled receptor related protein; GPCR related protein;
KW   NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW   immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW   antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW   diabetes; immune disorder; AIDS; obesity; asthma;
KW   haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW   infection; multiple sclerosis; cancer-associated cachexia;
KW   wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW   cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO200299116-A2.
XX
PD   12-DEC-2002.
XX
PF   04-JUN-2002; 2002WO-US017428.
XX
PR   04-JUN-2001; 2001US-0295607P.
PR   04-JUN-2001; 2001US-0295661P.
PR   06-JUN-2001; 2001US-0296404P.
PR   06-JUN-2001; 2001US-0296418P.
PR   14-JUN-2001; 2001US-0298285P.
PR   15-JUN-2001; 2001US-0298556P.
PR   21-JUN-2001; 2001US-0299949P.
PR   26-JUN-2001; 2001US-0300883P.
PR   28-JUN-2001; 2001US-0301550P.
PR   13-AUG-2001; 2001US-0311972P.
PR   27-AUG-2001; 2001US-0315071P.
PR   29-AUG-2001; 2001US-0315660P.
PR   14-SEP-2001; 2001US-0322293P.
PR   17-SEP-2001; 2001US-0322706P.
PR   14-DEC-2001; 2001US-0341186P.
PR   28-FEB-2002; 2002US-0361189P.
PR   12-MAR-2002; 2002US-0363673P.
PR   12-MAR-2002; 2002US-0363676P.
PR   03-JUN-2002; 2002US-00363676.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Anderson DW,  Baumgartner JC,  Boldog FL,  Casman SJ,  Edinger SR;
PI   Gangolli EA,  Gerlach VL,  Gorman L,  Guo X,  Hjalt T,  Kekuda R,  Li L;
PI   Macdougall JR,  Malyankar UM,  Millet I,  Padigaru M,  Patturajan M;
PI   Pena CEA,  Rastelli L,  Shimkets RA,  Stone DJ,  Spytek KA,  Vernet CAM;
PI   Voss EZ,  Zerhusen BD;
XX
DR   WPI; 2003-140627/13.
DR   N-PSDB; ACD03653.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 1; Page 147; 332pp; English.
XX
CC   The invention describes an isolated polypeptide (I) comprising any of 27
CC   118-961 residue amino acid sequences, given in the specification, a
CC   mature form of them, a sequence that is at least 95 % identical to them,
CC   or a sequence having one or more conservative substitutions in them. The
CC   polypeptide is useful in manufacturing a medicament for treating a
CC   syndrome associated with a human disease selected from a pathology
CC   associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC   and antibodies are useful in treating or preventing NOVX-associated
CC   disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC   disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC   disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC   associated cachexia, and other wasting disorders associated with chronic
CC   diseases. The nucleic acids and polypeptides may also be used as targets
CC   for the identification of small molecules that modulate or inhibit e.g.
CC   neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC   wound healing and angiogenesis, in gene therapy, in generation of
CC   antibodies that bind immunospecifically to NOVX substances for use in
CC   therapeutic or diagnostic methods. The nucleic acids are further used as
CC   hybridisation probes, in chromosome mapping, tissue typing, preventive
CC   medicine, and pharmacogenomics. The polypeptides are also useful as
CC   vaccines. This is the amino acid sequence of a novel human G-protein
CC   coupled receptor related protein NOV
XX
SQ   Sequence 390 AA;

  Query Match             90.0%;  Score 618;  DB 6;  Length 390;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 275

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 331


RESULT 7
ABU99143
ID   ABU99143 standard; protein; 398 AA.
XX
AC   ABU99143;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Novel human GPCR related protein NOV12a.
XX
KW   Human; G-protein coupled receptor related protein; GPCR related protein;
KW   NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW   immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW   antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
```

KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03647.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 143; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 398 AA;

Query Match 90.0%; Score 618; DB 6; Length 398;
Best Local Similarity 100.0%; Pred. No. 2.8e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
Db 224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
Db 284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 339

RESULT 8
ADE76864
ID ADE76864 standard; protein; 427 AA.
XX
AC ADE76864;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein expressed in a liver disorder #9.
XX
KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003108871-A1.
XX
PD 12-JUN-2003.
XX
PF 30-JUL-2001; 2001US-00919039.
XX
PR 28-JUL-2000; 2000US-0222113P.
XX
PA (KASE/) KASER M R.
XX
PI Kaser MR;
XX
DR WPI; 2004-031227/03.
DR N-PSDB; ADE76863.
XX
PT Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX
PS Claim 1; SEQ ID NO 29; 41pp; English.
XX
CC The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a

```
CC   liver disorder.
XX
SQ   Sequence 427 AA;

  Query Match             90.0%;  Score 618;  DB 8;  Length 427;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368


RESULT 9
ABU99144
ID   ABU99144 standard; protein; 615 AA.
XX
AC   ABU99144;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Novel human GPCR related protein NOV12b.
XX
KW   Human; G-protein coupled receptor related protein; GPCR related protein;
KW   NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW   immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW   antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW   diabetes; immune disorder; AIDS; obesity; asthma;
KW   haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW   infection; multiple sclerosis; cancer-associated cachexia;
KW   wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW   cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO200299116-A2.
XX
PD   12-DEC-2002.
XX
PF   04-JUN-2002; 2002WO-US017428.
XX
PR   04-JUN-2001; 2001US-0295607P.
PR   04-JUN-2001; 2001US-0295661P.
PR   06-JUN-2001; 2001US-0296404P.
PR   06-JUN-2001; 2001US-0296418P.
PR   14-JUN-2001; 2001US-0298285P.
PR   15-JUN-2001; 2001US-0298556P.
PR   21-JUN-2001; 2001US-0299949P.
PR   26-JUN-2001; 2001US-0300883P.
PR   28-JUN-2001; 2001US-0301550P.
PR   13-AUG-2001; 2001US-0311972P.
PR   27-AUG-2001; 2001US-0315071P.
PR   29-AUG-2001; 2001US-0315660P.
PR   14-SEP-2001; 2001US-0322293P.
PR   17-SEP-2001; 2001US-0322706P.
PR   14-DEC-2001; 2001US-0341186P.
PR   28-FEB-2002; 2002US-0361189P.
PR   12-MAR-2002; 2002US-0363673P.
PR   12-MAR-2002; 2002US-0363676P.
PR   03-JUN-2002; 2002US-00363676.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Anderson DW,  Baumgartner JC,  Boldog FL,  Casman SJ,  Edinger SR;
PI   Gangolli EA,  Gerlach VL,  Gorman L,  Guo X,  Hjalt T,  Kekuda R,  Li L;
PI   Macdougall JR,  Malyankar UM,  Millet I,  Padigaru M,  Patturajan M;
PI   Pena CEA,  Rastelli L,  Shimkets RA,  Stone DJ,  Spytek KA,  Vernet CAM;
PI   Voss EZ,  Zerhusen BD;
XX
DR   WPI; 2003-140627/13.
DR   N-PSDB; ACD03648.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT   atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 1; Page 144; 332pp; English.
XX
CC   The invention describes an isolated polypeptide (I) comprising any of 27
CC   118-961 residue amino acid sequences, given in the specification, a
CC   mature form of them, a sequence that is at least 95 % identical to them,
CC   or a sequence having one or more conservative substitutions in them. The
CC   polypeptide is useful in manufacturing a medicament for treating a
CC   syndrome associated with a human disease selected from a pathology
CC   associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC   and antibodies are useful in treating or preventing NOVX-associated
CC   disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC   disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC   disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC   associated cachexia, and other wasting disorders associated with chronic
CC   diseases. The nucleic acids and polypeptides may also be used as targets
CC   for the identification of small molecules that modulate or inhibit e.g.
CC   neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC   wound healing and angiogenesis, in gene therapy, in generation of
CC   antibodies that bind immunospecifically to NOVX substances for use in
CC   therapeutic or diagnostic methods. The nucleic acids are further used as
CC   hybridisation probes, in chromosome mapping, tissue typing, preventive
CC   medicine, and pharmacogenomics. The polypeptides are also useful as
CC   vaccines. This is the amino acid sequence of a novel human G-protein
CC   coupled receptor related protein NOV
XX
SQ   Sequence 615 AA;

  Query Match             90.0%;  Score 618;  DB 6;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 339


RESULT 10
ABB78707
ID   ABB78707 standard; protein; 626 AA.
XX
AC   ABB78707;
XX
DT   18-JUL-2002  (first entry)
XX
DE   Human high molecular weight kininogen (HK) mature protein SEQ ID NO:1.
XX
KW   Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW   D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumour;
KW   antiatherosclerotic; vasotropic; vulnerary; tranquilliser; thrombolytic;
KW   ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW   antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          384. .508
FT                   /label= D5_domain
XX
PN   WO200214369-A2.
```

XX
PD 21-FEB-2002.
XX
PF 24-JUL-2001; 2001WO-US023185.
XX
PR 24-JUL-2000; 2000US-0220194P.
XX
PA (ATTE-) ATTENUON LLC.
XX
PI Mazar AP, Juarez JC;
XX
DR WPI; 2002-393611/42.
XX
PT Novel human kininogen D5 domain polypeptides useful for treating
PT conditions associated with endothelial cell migration, proliferation,
PT invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
PT hyperplasia.
XX
PS Disclosure; Page 13; 84pp; English.
XX
CC The present invention describes an isolated polypeptide (I) that
CC corresponds to the D5 domain of human kininogen, or biologically active
CC peptide fragment, homologue or functional derivative, and which: (a)
CC inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
CC cells (EC); (c) activates signalling pathways leading to the introduction
CC of apoptosis in EC; and/or (d) inhibits the signalling pathway required
CC for maintenance of EC viability. (I) has cytostatic, antitumour,
CC antiatherosclerotic, vasotropic, vulnerary, tranquilliser, thrombolytic,
CC ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
CC antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
CC specific for an epitope of (I) is useful for inhibiting tumour growth or
CC angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
CC or trimeric fusion polypeptide (III) can be used for inhibiting EC
CC migration, proliferation, invasion, or angiogenesis, or for inducing EC
CC apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
CC comprising (I), (II), or (III), can be used for treating a subject having
CC a disease or condition associated with undesired EC migration,
CC proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
CC for isolating a D5 domain binding molecule from a complex mixture and for
CC isolating or enriching cells expressing D5 domain binding sites from a
CC cell mixture. The present sequence represents the mature human high
CC molecular weight kininogen (HK) protein, which is given in the
CC exemplification of the present invention
XX
SQ Sequence 626 AA;

Query Match 90.0%; Score 618; DB 5; Length 626;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 294

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 350
```

RESULT 11
ABG21101
ID ABG21101 standard; protein; 644 AA.
XX
AC ABG21101;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21092.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.

XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85288.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 51460; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 644 AA;

Query Match 90.0%; Score 618; DB 4; Length 644;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368
```

RESULT 12
ABB78710
ID ABB78710 standard; protein; 644 AA.
XX
AC ABB78710;
XX
DT 18-JUL-2002 (first entry)
XX
DE Human high molecular weight kininogen (HK) protein.
XX
KW Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumour;

```
KW   antiatherosclerotic; vasotropic; vulnerary; tranquilliser; thrombolytic;
KW   ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW   antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= signal
FT   Protein         19. .644
FT                   /label= mature_human_high_molecular_weight_kininogen
FT   Disulfide-bond  28. .614
FT   Disulfide-bond  83. .94
FT   Disulfide-bond  107. .126
FT   Disulfide-bond  142. .145
FT   Disulfide-bond  206. .218
FT   Disulfide-bond  229. .248
FT   Disulfide-bond  264. .267
FT   Disulfide-bond  328. .340
FT   Disulfide-bond  351. .370
FT   Domain          402. .526
FT                   /label= D5_domain
XX
PN   WO200214369-A2.
XX
PD   21-FEB-2002.
XX
PF   24-JUL-2001; 2001WO-US023185.
XX
PR   24-JUL-2000; 2000US-0220194P.
XX
PA   (ATTE-) ATTENUON LLC.
XX
PI   Mazar AP, Juarez JC;
XX
DR   WPI; 2002-393611/42.
XX
PT   Novel human kininogen D5 domain polypeptides useful for treating
PT   conditions associated with endothelial cell migration, proliferation,
PT   invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
PT   hyperplasia.
XX
PS   Disclosure; Fig 1B-E; 84pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) that
CC   corresponds to the D5 domain of human kininogen, or biologically active
CC   peptide fragment, homologue or functional derivative, and which: (a)
CC   inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
CC   cells (EC); (c) activates signalling pathways leading to the introduction
CC   of apoptosis in EC; and/or (d) inhibits the signalling pathway required
CC   for maintenance of EC viability. (I) has cytostatic, antitumour,
CC   antiatherosclerotic, vasotropic, vulnerary, tranquilliser, thrombolytic,
CC   ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
CC   antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
CC   specific for an epitope of (I) is useful for inhibiting tumour growth or
CC   angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
CC   or trimeric fusion polypeptide (III) can be used for inhibiting EC
CC   migration, proliferation, invasion, or angiogenesis, or for inducing EC
CC   apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
CC   comprising (I), (II), or (III), can be used for treating a subject having
CC   a disease or condition associated with undesired EC migration,
CC   proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
CC   for isolating a D5 domain binding molecule from a complex mixture and for
CC   isolating or enriching cells expressing D5 domain binding sites from a
CC   cell mixture. The present sequence represents the human high molecular
CC   weight kininogen (HK) protein, which is given in the exemplification of
CC   the present invention
XX
SQ   Sequence 644 AA;

  Query Match             90.0%;  Score 618;  DB 5;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-62;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368


RESULT 13
ABU99150
ID   ABU99150 standard; protein; 644 AA.
XX
AC   ABU99150;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Novel human GPCR related protein NOV12h.
XX
KW   Human; G-protein coupled receptor related protein; GPCR related protein;
KW   NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW   immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW   antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW   diabetes; immune disorder; AIDS; obesity; asthma;
KW   haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW   infection; multiple sclerosis; cancer-associated cachexia;
KW   wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW   cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO200299116-A2.
XX
PD   12-DEC-2002.
XX
PF   04-JUN-2002; 2002WO-US017428.
XX
PR   04-JUN-2001; 2001US-0295607P.
PR   04-JUN-2001; 2001US-0295661P.
PR   06-JUN-2001; 2001US-0296404P.
PR   06-JUN-2001; 2001US-0296418P.
PR   14-JUN-2001; 2001US-0298285P.
PR   15-JUN-2001; 2001US-0298556P.
PR   21-JUN-2001; 2001US-0299949P.
PR   26-JUN-2001; 2001US-0300883P.
PR   28-JUN-2001; 2001US-0301550P.
PR   13-AUG-2001; 2001US-0311972P.
PR   27-AUG-2001; 2001US-0315071P.
PR   29-AUG-2001; 2001US-0315660P.
PR   14-SEP-2001; 2001US-0322293P.
PR   17-SEP-2001; 2001US-0322706P.
PR   14-DEC-2001; 2001US-0341186P.
PR   28-FEB-2002; 2002US-0361189P.
PR   12-MAR-2002; 2002US-0363673P.
PR   12-MAR-2002; 2002US-0363676P.
PR   03-JUN-2002; 2002US-00363676.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI   Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI   Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI   Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI   Voss EZ, Zerhusen BD;
XX
DR   WPI; 2003-140627/13.
DR   N-PSDB; ACD03654.
XX
PT   New NOVX polypeptides and nucleic acids, useful for preventing or
PT   treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
```

PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 148; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 644 AA;

Query Match 90.0%; Score 618; DB 6; Length 644;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368

RESULT 14
ABU99145
ID ABU99145 standard; protein; 644 AA.
XX
AC ABU99145;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human GPCR related protein NOV12c.
XX
KW Human; G-protein coupled receptor related protein; GPCR related protein;
KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03649.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 144-145; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 644 AA;

Query Match 90.0%; Score 618; DB 6; Length 644;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368


RESULT 15
AAB37447
ID   AAB37447 standard; protein; 122 AA.
XX
AC   AAB37447;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human kininogen D3.
XX
KW   Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS   Homo sapiens.
XX
PN   WO200064945-A1.
XX
PD   02-NOV-2000.
XX
PF   20-APR-2000; 2000WO-GB001571.
XX
PR   22-APR-1999;   99GB-00009133.
XX
PA   (BABR-) BABRAHAM INST.
XX
PI   Abrahamson M,  Barrett AJ;
XX
DR   WPI; 2000-687316/67.
XX
PT   Inhibition of mammalian legumain or legumain-related endopeptidase by
PT   cystatin involves interaction with second papain-non-reactive site of
PT   cystatin.
XX
PS   Disclosure; Fig 4; 45pp; English.
XX
CC   The present invention relates to inhibition of the enzymatic activity of
CC   legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC   involves an interaction between legumain and a papain-non-reactive site
CC   of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC   performs a protein-processing function. The present sequence is human
CC   kininogen D3, which was used in the present invention. Kininogen is a
CC   type 3 cystatin
XX
SQ   Sequence 122 AA;

  Query Match             85.3%;  Score 586;  DB 3;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-59;
  Matches  110;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 PPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYF 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYF 60

Qy         69 IDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 110


Search completed: September 24, 2004, 14:08:38
Job time : 52.308 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:07:01 ; Search time 14.732 Seconds
(without alignments)
445.051 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	558	81.2	117	1	US-08-193-114B-1	Sequence 1, Appli
2	556.5	81.0	117	5	PCT-US92-06809-1	Sequence 1, Appli
3	163.5	23.8	145	2	US-08-832-535-2	Sequence 2, Appli
4	163.5	23.8	145	3	US-09-019-485-2	Sequence 2, Appli
5	163.5	23.8	145	3	US-09-019-485-3	Sequence 3, Appli
6	163.5	23.8	145	3	US-09-431-480-9	Sequence 9, Appli
7	163.5	23.8	145	3	US-09-617-302-9	Sequence 9, Appli
8	163.5	23.8	145	4	US-09-528-436B-2	Sequence 2, Appli
9	163	23.7	178	2	US-08-791-522-1	Sequence 1, Appli
10	163	23.7	178	3	US-09-314-777-1	Sequence 1, Appli
11	138.5	20.2	121	4	US-09-775-932-14	Sequence 14, Appl
12	138.5	20.2	128	4	US-09-775-932-12	Sequence 12, Appl
13	138.5	20.2	149	2	US-08-461-030C-2	Sequence 2, Appli
14	138.5	20.2	149	3	US-08-744-138-2	Sequence 2, Appli
15	138.5	20.2	149	3	US-09-431-480-8	Sequence 8, Appli
16	138.5	20.2	149	3	US-09-431-480-10	Sequence 10, Appl
17	138.5	20.2	149	3	US-09-617-302-8	Sequence 8, Appli
18	138.5	20.2	149	3	US-09-617-302-10	Sequence 10, Appl
19	138.5	20.2	149	4	US-09-241-376-2	Sequence 2, Appli
20	138.5	20.2	149	4	US-09-940-497-2	Sequence 2, Appli
21	137.5	20.0	112	4	US-08-849-303-16	Sequence 16, Appl
22	136.5	19.9	118	4	US-09-775-932-24	Sequence 24, Appl
23	135.5	19.7	146	6	5432264-6	Patent No. 5432264
24	134	19.5	148	5	PCT-US95-07135-2	Sequence 2, Appli
25	132.5	19.3	120	4	US-09-775-932-2	Sequence 2, Appli
26	132.5	19.3	145	2	US-08-832-535-11	Sequence 11, Appl
27	132.5	19.3	146	2	US-08-791-522-3	Sequence 3, Appli
28	132.5	19.3	146	3	US-08-744-138-3	Sequence 3, Appli
29	132.5	19.3	146	3	US-09-019-485-4	Sequence 4, Appli
30	132.5	19.3	146	3	US-09-314-777-3	Sequence 3, Appli
31	132.5	19.3	146	3	US-09-431-480-6	Sequence 6, Appli
32	132.5	19.3	146	3	US-09-617-302-6	Sequence 6, Appli
33	132.5	19.3	146	4	US-09-241-376-3	Sequence 3, Appli
34	132.5	19.3	146	4	US-09-528-436B-3	Sequence 3, Appli
35	132.5	19.3	146	4	US-09-886-319A-47	Sequence 47, Appl
36	132.5	19.3	146	4	US-09-940-497-3	Sequence 3, Appli
37	132.5	19.3	146	4	US-09-976-594-37	Sequence 37, Appl
38	132.5	19.3	146	4	US-08-849-303-17	Sequence 17, Appl
39	132.5	19.3	146	5	PCT-US95-07135-9	Sequence 9, Appli
40	132	19.2	26	3	US-08-676-242-15	Sequence 15, Appl
41	131.5	19.1	382	4	US-09-599-360B-93	Sequence 93, Appl
42	130	18.9	127	4	US-08-849-303-19	Sequence 19, Appl
43	129.5	18.9	140	4	US-09-886-319A-46	Sequence 46, Appl
44	129.5	18.9	140	4	US-09-886-319A-48	Sequence 48, Appl
45	128	18.6	111	4	US-08-849-303-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
; GENERAL INFORMATION:
; APPLICANT: Schmaier, Alvin H.
; APPLICANT: Jiang, Yongping
; TITLE OF INVENTION: Modulation of Blood
; TITLE OF INVENTION: Pressure and Inhibition of Platelet Activation
; TITLE OF INVENTION: with Kininogen Fragment
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna &
; ADDRESSEE: Monaco, P.C.
; STREET: 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,114B
; FILING DATE: 9 February 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Application
; APPLICATION NUMBER: Serial No. 5472945 07/744,545
; FILING DATE: 13 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-137 CI1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5472945e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: peptide
; TOPOLOGY: linear
US-08-193-114B-1

Query Match 81.2%; Score 558; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.3e-55;
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          14 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 73
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 60

Qy          74 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               |||||||||||||||||||||||||||||||||||||||||||||
Db          61 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 105


RESULT 2
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of the
;      ADDRESSEE:  Commonwealth System of Higher Education
;      STREET:  406 University Services
;      STREET:  Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06809
;      FILING DATE:  19910813
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No.  744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06809-1

  Query Match             81.0%;  Score 556.5;  DB 5;  Length 117;
  Best Local Similarity   93.8%;  Pred. No. 3.3e-55;
  Matches  106;  Conservative    1;  Mismatches    1;  Indels    5;  Gaps    1;

Qy          14 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 73
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 60

Qy          74 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWECE 126
               |||||||||||||||||||||:||||||||||||||||||||||       ||
Db          61 RETTCSKESNEELTESCETKKLGESLDCNAEVYVVPWEKKIYPTV-----NCE 108


RESULT 3
US-08-832-535-2
; Sequence 2, Application US/08832535
; Patent No. 5919658
;  GENERAL INFORMATION:
;    APPLICANT:  NI, JIAN
;    APPLICANT:  LI, HAODONG
;    APPLICANT:  YU, GUO-LIANG
;    APPLICANT:  GENTZ, REINER L
;    TITLE OF INVENTION:  HUMAN CYSTATIN F
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/832,535
;      FILING DATE:  03-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIMBALL, PAUL C.
;      REGISTRATION NUMBER:  34,610
;      REFERENCE/DOCKET NUMBER:  PF265
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-832-535-2

  Query Match             23.8%;  Score 163.5;  DB 2;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy          11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
               :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                  ||| |  :  | : |:       | |:| | :||:||||         : |:|
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 4
US-09-019-485-2
; Sequence 2, Application US/09019485
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-2

  Query Match             23.8%;  Score 163.5;  DB 3;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy          11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
               :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                  ||| |  :  | : |:      | |:| | :||:|||          : |:|
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 5
US-09-019-485-3
; Sequence 3, Application US/09019485
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-3

  Query Match             23.8%;  Score 163.5;  DB 3;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy          11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
               :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                  ||| |  :  | : |:      | |:| | :||:|||          : |:|
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 6
US-09-431-480-9
; Sequence 9, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
;  APPLICANT: Holloway, James L.
;  APPLICANT: Feldhaus, Andrew
;  TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE: 98-72
;  CURRENT APPLICATION NUMBER: US/09/431,480
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: 60/109,217
;  EARLIER FILING DATE: 1998-11-20
;  EARLIER APPLICATION NUMBER: 60/156,382
;  EARLIER FILING DATE: 1999-09-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-431-480-9

  Query Match             23.8%;  Score 163.5;  DB 3;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy          11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
               :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                  ||| |  :  | : |:      | |:| | :||:|||          : |:|
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 7
US-09-617-302-9
; Sequence 9, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
;  APPLICANT:  Holloway, James L.
;  APPLICANT:  Feldhaus, Andrew
;  TITLE OF INVENTION:  TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE:  98-72 C1
;  CURRENT APPLICATION NUMBER: US/09/617,302
;  CURRENT FILING DATE:  2000-07-17
;  PRIOR APPLICATION NUMBER:  09/431,480
;  PRIOR FILING DATE:  1999-11-01
;  PRIOR APPLICATION NUMBER:  60/109,217
;  PRIOR FILING DATE:  1998-11-20
;  PRIOR APPLICATION NUMBER:  60/156,382
;  PRIOR FILING DATE:  1999-09-28
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-617-302-9

  Query Match             23.8%;  Score 163.5;  DB 3;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |  :  | : |:      | |:| | :||:||||         : |:|
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 8
US-09-528-436B-2
; Sequence 2, Application US/09528436B
; Patent No. 6576745
; GENERAL INFORMATION:
;  APPLICANT: Li, et al.
;  TITLE OF INVENTION: Human Cystatin F
;  FILE REFERENCE: PF265P1D1
;  CURRENT APPLICATION NUMBER: US/09/528,436B
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER: 09/019,485
;  PRIOR FILING DATE: 1998-01-29
;  PRIOR APPLICATION NUMBER: 08/832,535
;  PRIOR FILING DATE: 1999-04-03
;  PRIOR APPLICATION NUMBER: 60/014,795
;  PRIOR FILING DATE: 1996-04-03
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-528-436B-2

  Query Match             23.8%;  Score 163.5;  DB 4;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.4e-11;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |  :  | : |:      | |:| | :||:||||         : |:|
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 9
US-08-791-522-1
; Sequence 1, Application US/08791522
; Patent No. 5935817
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/791,522
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-08-791-522-1

  Query Match             23.7%;  Score 163;  DB 2;  Length 178;
  Best Local Similarity   34.0%;  Pred. No. 1.4e-10;
  Matches   35;  Conservative   20;  Mismatches   44;  Indels    4;  Gaps    2;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPW 110
                 ||| |  :  | : |:      | |:| | :||:||||
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW 155


RESULT 10
US-09-314-777-1
; Sequence 1, Application US/09314777
; Patent No. 6110686
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/314,777
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/791,522
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-09-314-777-1

  Query Match             23.7%;  Score 163;  DB 3;  Length 178;
  Best Local Similarity   34.0%;  Pred. No. 1.4e-10;
  Matches   35;  Conservative   20;  Mismatches   44;  Indels    4;  Gaps    2;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPW 110
                 ||| |  :  | : |:       | |:| | :||:|||
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW 155


RESULT 11
US-09-775-932-14
; Sequence 14, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 121
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-14

  Query Match             20.2%;  Score 138.5;  DB 4;  Length 121;
  Best Local Similarity   31.5%;  Pred. No. 4.9e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          2 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 61

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         62 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 108


RESULT 12
US-09-775-932-12
; Sequence 12, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-12

  Query Match             20.2%;  Score 138.5;  DB 4;  Length 128;
  Best Local Similarity   31.5%;  Pred. No. 5.2e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          9 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 68

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         69 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 115


RESULT 13
US-08-461-030C-2
; Sequence 2, Application US/08461030C
; Patent No. 5985601
;  GENERAL INFORMATION:
;    APPLICANT:  Ni, Jian
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Rosen, Craig A.
;    TITLE OF INVENTION:  Human Cystatin E
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,030C
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders, Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-301-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-461-030C-2

  Query Match             20.2%;  Score 138.5;  DB 2;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 6.4e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db         30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89
```

```
Qy          70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|      |    | |: || ||||:
Db          90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 14
US-08-744-138-2
; Sequence 2, Application US/08744138
; Patent No. 6011012
;  GENERAL INFORMATION:
;    APPLICANT:  Gentz, Reiner L.
;    APPLICANT:  Ni, Jian
;    APPLICANT:  Rosen, Craig A.
;    APPLICANT:  Yu, Guo-Liang
;    TITLE OF INVENTION:  Human Cystatin E
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/744,138
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF202P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301 309 8504
;      TELEFAX:  301 309 8512
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-744-138-2

  Query Match           20.2%;  Score 138.5;  DB 3;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 6.4e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
               | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy          70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|      |    | |: || ||||:
Db          90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 15
US-09-431-480-8
; Sequence 8, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
;  APPLICANT: Holloway, James L.
;  APPLICANT: Feldhaus, Andrew
;  TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE: 98-72
;  CURRENT APPLICATION NUMBER: US/09/431,480
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: 60/109,217
;  EARLIER FILING DATE: 1998-11-20
;  EARLIER APPLICATION NUMBER: 60/156,382
;  EARLIER FILING DATE: 1999-09-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-431-480-8

  Query Match           20.2%;  Score 138.5;  DB 3;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 6.4e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
               | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy          70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|      |    | |: || ||||:
Db          90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


Search completed: September 24, 2004, 14:11:37
Job time : 15.732 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:05:18 ; Search time 36.576 Seconds
(without alignments)
1095.549 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	383	55.7	140	6	Q7YRP6	Q7yrp6 sus scrofa
2	381	55.5	423	11	P70517	P70517 rattus norv
3	378	55.0	430	11	Q63581	Q63581 rattus norv
4	171.5	25.0	167	11	Q9QWL5	Q9qwl5 mus musculu
5	163.5	23.8	167	4	Q7Z4J8	Q7z4j8 homo sapien
6	152.5	22.2	462	13	Q7ZY91	Q7zy91 xenopus lae
7	152.5	22.2	462	13	Q7SYH2	Q7syh2 xenopus lae
8	152.5	22.2	465	13	Q801E5	Q801e5 xenopus lae
9	131.5	19.1	140	11	Q9EPX9	Q9epx9 mus musculu
10	123.5	18.0	455	13	Q800S8	Q800s8 brachydanio
11	119	17.3	388	11	Q8CB17	Q8cb17 mus musculu
12	117.5	17.1	464	13	Q801Z5	Q801z5 cyprinus ca
13	113.5	16.5	148	5	Q9NH95	Q9nh95 litomosoide
14	113	16.4	140	11	Q80Y72	Q80y72 mus musculu
15	111	16.2	146	11	Q8K397	Q8k397 mus musculu
16	111	16.2	149	11	Q9D1B1	Q9d1b1 mus musculu
17	108.5	15.8	112	13	Q98SR4	Q98sr4 acipenser s
18	108.5	15.8	112	13	Q98SR3	Q98sr3 acipenser s
19	107	15.6	139	5	Q9TYY2	Q9tyy2 caenorhabdi
20	106	15.4	300	13	Q801Z6	Q801z6 cyprinus ca
21	105	15.3	109	5	Q9TY65	Q9ty65 onchocerca
22	105	15.3	127	5	P90698	P90698 brugia mala
23	104.5	15.2	149	11	Q8VHC1	Q8vhc1 rattus norv
24	104.5	15.2	161	5	O16159	O16159 brugia mala
25	102.5	14.9	127	5	Q9U9A1	Q9u9a1 onchocerca
26	101	14.7	148	11	Q9JM84	Q9jm84 mus musculu
27	99	14.4	110	10	Q8SA65	Q8sa65 sandersonia
28	98.5	14.3	107	5	Q8T0Y2	Q8t0y2 sarcophaga
29	98.5	14.3	125	5	Q25620	Q25620 onchocerca
30	97.5	14.2	134	10	Q41825	Q41825 zea mays (m
31	95.5	13.9	143	5	O61973	O61973 caenorhabdi
32	94	13.7	122	5	O44396	O44396 haemonchus
33	93	13.5	138	4	Q8WXU6	Q8wxu6 homo sapien
34	92	13.4	157	5	Q17108	Q17108 acanthochei
35	90	13.1	139	11	Q8K5A3	Q8k5a3 rattus norv
36	88.5	12.9	92	10	Q9FXN6	Q9fxn6 arabidopsis
37	88.5	12.9	116	10	Q8RXS7	Q8rxs7 arabidopsis
38	88.5	12.9	124	10	Q41906	Q41906 arabidopsis
39	88.5	12.9	125	10	O22202	O22202 arabidopsis
40	88.5	12.9	134	10	P93627	P93627 zea mays (m
41	88.5	12.9	134	10	Q41897	Q41897 zea mays (m
42	88	12.8	199	10	Q39270	Q39270 brassica ca
43	87.5	12.7	141	11	Q9DAP1	Q9dap1 mus musculu
44	86.5	12.6	141	11	Q80ZN5	Q80zn5 mus musculu
45	86.5	12.6	349	6	O14502	O14502 cercopithec

ALIGNMENTS

RESULT 1
Q7YRP6
ID Q7YRP6 PRELIMINARY; PRT; 140 AA.
AC Q7YRP6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Low molecular weight kininogen (Fragment).
GN KNG.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Vonnahme K.A., Fernando S.C., Ross J.A., Ashworth M.D., DeSilva U.,
RA Malayer J.R., Geisert R.D.;
RT "Porcine Endometrial and Conceptus Expression of Kininogens and Plasma
RT Kallikrein in Cyclic and Pregnant Gilts.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY321363; AAP85260.1; -.
FT NON_TER 1 1
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15650 MW; 177837836603F777 CRC64;

Query Match 55.7%; Score 383; DB 6; Length 140;
Best Local Similarity 78.9%; Pred. No. 2.8e-30;
Matches 75; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

```
Qy         24 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 83
              :||:||| | |:| |||||||| ||||| |:|| ||||||||| | | |||||||:||
Db          1 DSPDLEEPLNHSIAKLNAENNAVFYFKIGPVEKATVQVVAGKKYSIVFTARETTCSKQSN 60

Qy         84 EELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              |||||||| || || | ||| ||||||||||||||
Db         61 EELTESCEIKKPGQILKCNASVYVVPWEKKIYPTV 95
```

RESULT 2

```
P70517
ID   P70517      PRELIMINARY;      PRT;   423 AA.
AC   P70517;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major acute phase alpha-1 protein precursor (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cole T.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85149311; PubMed=2579644;
RA   Cole T., Inglis A., Nagashima M., Schreiber G.;
RT   "Major acute-phase alpha(1)-protein in the rat: Structure, molecular
RT   cloning, and regulation of mRNA levels.";
RL   Biochem. Biophys. Res. Commun. 126:719-724(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
DR   EMBL; K02814; AAA41569.1; -.
DR   PIR; S68034; S68034.
DR   PIR; S68035; S68035.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Signal.
FT   NON_TER      1      1
FT   SIGNAL       <1     11       POTENTIAL.
FT   CHAIN        12     423      POTENTIAL.
FT   CHAIN        371    379      POTENTIAL.
SQ   SEQUENCE   423 AA;  46905 MW;  F9E8BD3198547949 CRC64;

  Query Match             55.5%;  Score 381;  DB 11;  Length 423;
  Best Local Similarity   62.1%;  Pred. No. 1.6e-29;
  Matches   72;  Conservative   14;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        245 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||  | |:|:|||| |||:|  |||  |||| ||||:||| ||: ||| |: |||
Db        305 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 360


RESULT 3
Q63581
ID   Q63581      PRELIMINARY;      PRT;   430 AA.
AC   Q63581;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Rat T-kininogen (T-KG).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90034172; PubMed=2806908;
RA   Anderson K.P., Croyle M.L., Lingrel J.B.;
RT   "Primary structure of a gene encoding rat T-kininogen.";
RL   Gene 81:119-128(1989).
DR   EMBL; M29090; AAA42251.1; -.
DR   EMBL; M29083; AAA42251.1; JOINED.
DR   EMBL; M29084; AAA42251.1; JOINED.
DR   EMBL; M29091; AAA42251.1; JOINED.
DR   EMBL; M29085; AAA42251.1; JOINED.
DR   EMBL; M29086; AAA42251.1; JOINED.
DR   EMBL; M29087; AAA42251.1; JOINED.
DR   EMBL; M29088; AAA42251.1; JOINED.
DR   EMBL; M29089; AAA42251.1; JOINED.
DR   PIR; S68034; S68034.
DR   PIR; S68035; S68035.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
SQ   SEQUENCE   430 AA;  47618 MW;  45508DEF4BDC978C CRC64;

  Query Match             55.0%;  Score 378;  DB 11;  Length 430;
  Best Local Similarity   62.1%;  Pred. No. 3.1e-29;
  Matches   72;  Conservative   13;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |       | | ||||:|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        252 GDDLFSLLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||  | |:|:|||| |||:|  |||  |||| ||||:||| ||: ||| |: |||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 4
Q9QWL5
ID   Q9QWL5      PRELIMINARY;      PRT;   167 AA.
AC   Q9QWL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Murine CMAP (CYSTATIN P) (LEUKOCYSTATIN).
GN   MURINE CMAP OR CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morita M., Arakawa H., Yoshiuchi N.;
RT   "A novel cystatin-like metastasis associated gene.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
```

```
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AB015224; BAA34940.1; -.
DR   EMBL; AK004420; BAB23298.1; -.
DR   HSSP; P01034; 1G96.
DR   MGD; MGI:1298217; Cst7.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   167 AA;  18847 MW;  61F776D8445095FE CRC64;
```

```
  Query Match             25.0%;  Score 171.5;  DB 11;  Length 167;
  Best Local Similarity   35.5%;  Pred. No. 3.4e-09;
  Matches   39;  Conservative   22;  Mismatches   42;  Indels    7;  Gaps    3;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              ||  :   : :   | |: | ||:| : :   |:: | |    |  | ||  :| || ||||
Db         50 KDLI---SSVKPGFPKTIETNNPGVLKAARHSVEKFNNCTNDIFLFKESHVSKALVQVVK 106

Qy         64 GKKYFIDFVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPW 110
              | ||  ::     ||| |  : :| ::|:       | ::| | :||:|:||
Db        107 GLKYMLEVKIGRTTCRKTMHHQL-DNCDFQTNPALKRTLYCYSEVWVIPW 155
```

```
RESULT 5
Q7Z4J8
ID   Q7Z4J8      PRELIMINARY;      PRT;   167 AA.
AC   Q7Z4J8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin F (Leukocystatin).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT009825; AAP88827.1; -.
SQ   SEQUENCE   167 AA;  18857 MW;  E339025A5BD60177 CRC64;
```

```
  Query Match             23.8%;  Score 163.5;  DB 4;  Length 167;
  Best Local Similarity   31.6%;  Pred. No. 2.1e-08;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |   :  | : |:       | |:| | :||:|||          : |:|
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 160
```

```
RESULT 6
Q7ZY91
ID   Q7ZY91      PRELIMINARY;      PRT;   462 AA.
AC   Q7ZY91;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to fetuin B.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043891; AAH43891.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
SQ   SEQUENCE   462 AA;  53185 MW;  D7BAD339961739FB CRC64;
```

```
  Query Match             22.2%;  Score 152.5;  DB 13;  Length 462;
  Best Local Similarity   38.8%;  Pred. No. 8.4e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 65
              |  |    || | || : |:  |:| |     | : | ::| | |||||:::: | | | |
Db        142 PGVILSTCP-DCPTANEEITPTITETADTLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 200

Qy         66 KYFIDFVARETTCSKESNEELTESC 90
              ||| |  :|| | |      :  :|
Db        201 SYFIQFTIKETDCMKTQENVVLSNC 225
```

```
RESULT 7
Q7SYH2
ID   Q7SYH2      PRELIMINARY;      PRT;   462 AA.
AC   Q7SYH2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin domain fetuin-like protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ventral midgut;
RA   Costa R.M.B., Mason J., Lee M., Amaya E., Zorn A.M.;
RT   "Novel gene expression domains reveal early patterning of the Xenopus
RT   endoderm.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY260732; AAP82289.1; -.
SQ   SEQUENCE   462 AA;  53186 MW;  796F92774CC27721 CRC64;
```

```
  Query Match             22.2%;  Score 152.5;  DB 13;  Length 462;
  Best Local Similarity   38.8%;  Pred. No. 8.4e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 65
              |  |    || | || : |:  |:| |     | : | ::| | |||||:::: | | | |
Db        142 PGVILSTCP-DCPTANEEITPTITETADTLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 200

Qy         66 KYFIDFVARETTCSKESNEELTESC 90
              ||| |  :|| | |      :  :|
Db        201 SYFIQFTIKETDCMKTQENVVLSNC 225
```

```
RESULT 8
Q801E5
ID   Q801E5      PRELIMINARY;      PRT;   465 AA.
AC   Q801E5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical histidine-rich protein (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
```

```
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22480013; PubMed=12591597;
RA   Chen Y., Jurgens K., Hollemann T., Claussen M., Ramadori G.,
RA   Pieler T.;
RT   "Cell-autonomous and signal-dependent expression of liver and
RT   intestine marker genes in pluripotent precursor cells from Xenopus
RT   embryos.";
RL   Mech. Dev. 120:277-288(2003).
DR   EMBL; AY188284; AAO31610.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   465 AA;  53528 MW;  0B403AB4P78BBFD4 CRC64;

  Query Match             22.2%;  Score 152.5;  DB 13;  Length 465;
  Best Local Similarity   38.8%;  Pred. No. 8.5e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 65
              |  |   || | || : |:  |:| |     | : | ::| | |||||:::: | | | |
Db        145 PGVILSTCP-DCPTANEEITPTITETAETLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 203

Qy         66 KYFIDFVARETTCSKESNEELTESC 90
                ||| |  :|| | |     :  :|
Db        204 SYFIQFTIKETDCMKTQENVVLSNC 228


RESULT 9
Q9EPX9
ID   Q9EPX9      PRELIMINARY;      PRT;   140 AA.
AC   Q9EPX9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cystatin C.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=21010502; PubMed=11144350;
RA   Taupin P.J., Ray J., Fischer W.H., Suhr S.T., Hakansson K., Grubb A.,
RA   Gage F.H.;
RT   "FGF-2-Responsive neural stem cell proliferation requires CCg, a novel
RT   autocrine/paracrine cofactor.";
RL   Neuron 28:385-397(2000).
DR   EMBL; AF311741; AAG40283.1; -.
DR   HSSP; P01034; 1G96.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
FT   CHAIN        21    140       CYSTATIN C.
FT   VARIANT      16     16       A -> G.
FT   VARIANT      84     84       L -> F.
SQ   SEQUENCE   140 AA;  15517 MW;  3A563406DD58D785 CRC64;

  Query Match             19.1%;  Score 131.5;  DB 11;  Length 140;
  Best Local Similarity   27.8%;  Pred. No. 2.6e-05;
  Matches   32;  Conservative   26;  Mismatches   48;  Indels    9;  Gaps    4;

Qy         15 VGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVAR 74
              :| | :   |   :   |   ::: |  :|  :: :   | :|| |:||| ||:|
Db         30 LGAPEEADANEEGVRRALDFAVSEYNKGSNDAYHSRAIQVVRARKQLVAGVNYFLDVEMG 89

Qy         75 ETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWECE 126
                 |||:| |   ||: |   :   | :   |: ::| |||:      | :: :: |:
Db         90 RTTCTK-SQTNLTD-CPFHDQPHLMRKALCSFQIYSVPWK----GTHSLTNFSCK 138


RESULT 10
Q800S8
ID   Q800S8      PRELIMINARY;      PRT;   455 AA.
AC   Q800S8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin-A.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Jia F.;
RT   "Danio rerio fetuin-A.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY217758; AAO61483.1; -.
DR   GO; GO:0005874; C:microtubule; IEA.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0007018; P:microtubule-based movement; IEA.
DR   InterPro; IPR002453; Beta_tubulin.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
SQ   SEQUENCE   455 AA;  50627 MW;  D822872926BA2ACB CRC64;

  Query Match             18.0%;  Score 123.5;  DB 13;  Length 455;
  Best Local Similarity   26.7%;  Pred. No. 0.00063;
  Matches   32;  Conservative   23;  Mismatches   46;  Indels   19;  Gaps    4;

Qy          2 SGKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQ- 60
              | :| |     | |  ||  :| : |:  |::   : | | ::|   |||:  | :   |
Db        134 SHEDLV----KKCPDCPGLLPLHEPKALESVNAALAKFNKQSNHKSYFKLMEVGRISSQW 189

Qy         61 VVAGKKYFIDFVARETTCSKESNEELTES------------CETKKLG-QSLDCNAEVY 106
              :  |: ||  |   || |:|:   :  |:             |:: |:| :  :   |:|
Db        190 MPMGQSYFTQFAIMETNCTKKDAPQNPEACKALCGDQATYGFCKSSKVGSEEPEVECEIY 249


RESULT 11
Q8CB17
ID   Q8CB17      PRELIMINARY;      PRT;   388 AA.
AC   Q8CB17;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin beta.
GN   FETUB.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Vagina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK037043; BAC29682.1; -.
DR   MGD; MGI:1890221; Fetub.
```

```
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR001363; Fetuin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
SQ   SEQUENCE   388 AA;  42742 MW;  78CFAD73A8D8DC22 CRC64;

  Query Match             17.3%;  Score 119;  DB 11;  Length 388;
  Best Local Similarity   25.9%;  Pred. No. 0.0015;
  Matches   30;  Conservative   28;  Mismatches   52;  Indels    6;  Gaps    3;

Qy         14 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 73
              |  ||  |  ::|   |  | :: | |::: :  | ::  | ||  | |:| |:::::
Db        154 CPDCPSPIDLSNPSALEAATESLAKFNSKSPSKKY-ELVKVTKAMNQWVSGPAYYVEYLI 212

Qy         74 RETTCSKESNEELTESCETKKLG--QSLDCNAEVYVVPWEKKIYPTVTVNHWECEF 127
              :|  |:|       :  ::: :|  |     : :  ||  : :  :|||    |||
Db        213 KEAPCTKSQASCSLQHSDSEPVGICQGSTVQSSLRHVPLIQPVEKSVTVT---CEF 265


RESULT 12
Q801Z5
ID   Q801Z5      PRELIMINARY;      PRT;   464 AA.
AC   Q801Z5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin long form.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tsai P.-L., Chang G.-D., Huang C.-J.;
RT   "Purification and cloning of carp fetuin.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY225965; AAO74862.1; -.
DR   GO; GO:0005874; C:microtubule; IEA.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0007018; P:microtubule-based movement; IEA.
DR   InterPro; IPR002453; Beta_tubulin.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
SQ   SEQUENCE   464 AA;  51698 MW;  7A54F71E44050895 CRC64;

  Query Match             17.1%;  Score 117.5;  DB 13;  Length 464;
  Best Local Similarity   24.8%;  Pred. No. 0.0025;
  Matches   32;  Conservative   23;  Mismatches   59;  Indels   15;  Gaps    4;

Qy         14 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQ-VVAGKKYFIDFV 72
              |  ||  :| : |:  |::   : | | |::   |||:  | :   | : :|: :|  |
Db        142 CPDCPGLLPLHDPKGLESVKTALQKFNKESDHKSYFKLMEVGRISTQWMFSGQSFFSQFA 201

Qy         73 ARETTCSKESNEELTESC-----ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNH-W--- 123
                || |: :   :  |||     |  : |         :     | :||    : | |
Db        202 IMETNCTNKEAPQNEESCKALCGEKARYGFCKSTKVGIEEPEVECEIYEAKNITHPWKHP 261

Qy        124 -----ECEF 127
                   :|:|
Db        262 AQSRRDCKF 270


RESULT 13
Q9NH95
ID   Q9NH95      PRELIMINARY;      PRT;   148 AA.
AC   Q9NH95;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ls-cystatin.
OS   Litomosoides sigmodontis.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC   Onchocercidae; Litomosoides.
OX   NCBI_TaxID=42156;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
RT   "Characterization and immunological properties of a cystein protease
RT   inhibitor of the filarial parasite Litomosoides sigmodontis.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF229173; AAF35896.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
FT   CHAIN        25    148       LS-CYSTATIN.
SQ   SEQUENCE   148 AA;  16686 MW;  2950AA89CA5339C9 CRC64;

  Query Match             16.5%;  Score 113.5;  DB 5;  Length 148;
  Best Local Similarity   34.1%;  Pred. No. 0.0017;
  Matches   30;  Conservative   16;  Mismatches   37;  Indels    5;  Gaps    3;

Qy         27 ELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESNEEL 86
              |::| |   :||:| ::|  ::     | |   ||||| || ::     : | | |||::
Db         49 EIQEMLPSILTKVNQQSNDAYHLMPIKVLKVSSQVVAGMKYKMEIQVARSDCKKSSNEKI 108

Qy         87 -TESCETKKLGQSLD--CNAEVYVVPWE 111
               ::|  |||     |     ||:  ||
Db        109 DLKTC--KKLEGHPDQIITLEVWEKAWE 134


RESULT 14
Q80Y72
ID   Q80Y72      PRELIMINARY;      PRT;   140 AA.
AC   Q80Y72;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin-like 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testicle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testicle;
RA   Strausberg R.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC048646; AAH48646.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C/M.
DR   Pfam; PF00031; cystatin; 1.
DR   ProDom; PD001231; Cystatin_C/M; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   140 AA;  16199 MW;  32633E99C4697DA0 CRC64;

  Query Match             16.4%;  Score 113;  DB 11;  Length 140;
  Best Local Similarity   29.4%;  Pred. No. 0.0018;
  Matches   25;  Conservative   18;  Mismatches   38;  Indels    4;  Gaps    2;

Qy         28 LEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESNEELT 87
              :  ||    |   | :| |: :::  : : ::|:  | :| :       | | |  ||
Db         44 INSTLHFFIRSYNNASNDTYLYQVQKLIQGQMQLTTGVEYLVTVKIGRTKCKK--NETKK 101

Qy         88 ESC--ETKKLGQSLDCNAEVYVVPW 110
                ||  :: || :|| | : :| |||
Db        102 ASCPLQSSKLKKSLICKSLIYSVPW 126


RESULT 15
Q8K397
ID   Q8K397      PRELIMINARY;      PRT;   146 AA.
AC   Q8K397;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RIKEN cDNA 1110017E11 gene (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC027680; AAH27680.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   146 AA;  16380 MW;  9D77BB9A6063A5C4 CRC64;

  Query Match             16.2%;  Score 111;  DB 11;  Length 146;
  Best Local Similarity   30.5%;  Pred. No. 0.003;
  Matches   29;  Conservative   18;  Mismatches   42;  Indels    6;  Gaps    4;

Qy         22 PTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKE 81
              ||: | :::     :   |  ::: :||:   |  |: |:||| ||::     | | |
Db         40 PTD-PRVQKAAQAAVASYNMGSDSLYYFRDTKVIDAKYQLVAGIKYYLTLDIESTECRKT 98

Qy         82 --SNEEL-TESCETKKLGQ--SLDCNAEVYVVPWE 111
                | | :   :|     ||   | || |:  |||:
Db         99 RVSGEHMDLTTCPLAAGGQQEKLRCNFELLEVPWK 133


Search completed: September 24, 2004, 14:10:18
Job time : 37.576 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:04:32 ; Search time 8.636 Seconds
(without alignments)
765.738 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	618	90.0	644	1	KNG_HUMAN	P01042 homo sapien
2	440	64.0	436	1	KNL1_BOVIN	P01046 bos taurus
3	440	64.0	621	1	KNH1_BOVIN	P01044 bos taurus
4	413	60.1	434	1	KNL2_BOVIN	P01047 bos taurus
5	413	60.1	619	1	KNH2_BOVIN	P01045 bos taurus
6	413	60.1	661	1	KNG_MOUSE	O08677 mus musculu
7	410	59.7	639	1	KNG_RAT	P08934 rattus norv
8	388	56.5	430	1	KNT2_RAT	P08932 rattus norv
9	380	55.3	430	1	KNT1_RAT	P01048 rattus norv
10	171.5	25.0	144	1	CYTF_MOUSE	O89098 mus musculu
11	163.5	23.8	145	1	CYTF_HUMAN	O76096 homo sapien
12	138.5	20.2	146	1	CYTC_MACMU	O19092 macaca mula
13	138.5	20.2	149	1	CYTM_HUMAN	Q15828 homo sapien
14	137.5	20.0	148	1	CYTC_BOVIN	P01035 bos taurus
15	135	19.7	378	1	FETB_RAT	Q9qx79 rattus norv
16	132.5	19.3	146	1	CYTC_HUMAN	P01034 homo sapien
17	132.5	19.3	146	1	CYTC_SAISC	O19093 saimiri sci
18	131.5	19.1	382	1	FETB_HUMAN	Q9ugm5 homo sapien
19	130	18.9	127	1	CYTC_RAT	P14841 rattus norv
20	129.5	18.9	140	1	CYTC_MOUSE	P21460 mus musculu
21	128	18.6	111	1	CYT_BITAR	P08935 bitis ariet
22	124.5	18.1	141	1	CYTT_HUMAN	P09228 homo sapien
23	124.5	18.1	148	1	CYTC_RABIT	O97862 oryctolagus
24	122.5	17.8	116	1	CYT_COTJA	P81061 coturnix co
25	119	17.3	388	1	FETB_MOUSE	Q9qxc1 mus musculu
26	118.5	17.2	139	1	CYT_CHICK	P01038 gallus gall
27	113	16.4	141	1	CYTS_RAT	P19313 rattus norv
28	109.5	15.9	141	1	CYTN_HUMAN	P01037 homo sapien
29	108.5	15.8	141	1	CYTS_HUMAN	P01036 homo sapien
30	107	15.6	130	1	CYT_ONCKE	Q98967 oncorhynchu
31	105.5	15.4	162	1	CYTX_ONCVO	P22085 onchocerca
32	105	15.3	130	1	CYT_ONCMY	Q91195 oncorhynchu
33	104	15.1	129	1	CYT_CYPCA	P35481 cyprinus ca
34	102	14.8	122	1	CYTA_SARPE	P31727 sarcophaga
35	101	14.7	139	1	CS11_MOUSE	Q9d269 mus musculu
36	95	13.8	165	1	CSTL_HUMAN	Q9h114 homo sapien
37	94.5	13.8	142	1	CST8_MOUSE	P32766 mus musculu
38	94.5	13.8	345	1	A2HS_MOUSE	P29699 mus musculu
39	94	13.7	142	1	CST8_HUMAN	O60676 homo sapien
40	93	13.5	99	1	CYT_NAJAT	P81714 naja atra (
41	93	13.5	137	1	CS11_HUMAN	Q9h112 homo sapien
42	93	13.5	142	1	CYTD_HUMAN	P28325 homo sapien
43	91.5	13.3	135	1	CYT1_MAIZE	P31726 zea mays (m
44	91.5	13.3	352	1	A2HS_RAT	P24090 rattus norv
45	88	12.8	367	1	A2HS_HUMAN	P02765 homo sapien

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

```
RL   Seikagaku 56:808-808(1984).
RN   [7]
RP   CARBOHYDRATE-LINKAGE SITE ASN-294.
RX   MEDLINE=22660472; PubMed=12754519;
RA   Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT   "Identification and quantification of N-linked glycoproteins using
RT   hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL   Nat. Biotechnol. 21:660-666(2003).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
CC       clotting.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=P01042-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=P01042-2; Sequence=VSP_001261, VSP_001262;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A01280; KGHUL1.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   Genew; HGNC:6383; KNG.
DR   MIM; 228960; -.
DR   GO; GO:0007596; P:blood coagulation; NAS.
DR   GO; GO:0030146; P:diuresis; NAS.
DR   GO; GO:0006954; P:inflammatory response; NAS.
DR   GO; GO:0030147; P:natriuresis; NAS.
DR   GO; GO:0006939; P:smooth muscle contraction; NAS.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Pyrrolidone carboxylic acid.
FT   SIGNAL       1    18
FT   CHAIN       19   644       KININOGEN.
FT   CHAIN       19   380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381   389       BRADYKININ.
FT   CHAIN      390   644       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19   136       CYSTATIN-LIKE 1.
FT   DOMAIN     137   258       CYSTATIN-LIKE 2.
FT   DOMAIN     259   380       CYSTATIN-LIKE 3.
FT   DOMAIN     420   510       HIS-RICH
FT                              (ASSOCIATED WITH CLOTTING ACTIVITY).
FT   REPEAT     420   449
FT   REPEAT     450   479
FT   REPEAT     480   510
FT   MOD_RES     19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID    28   614       INTERCHAIN.
FT   DISULFID    83    94
FT   DISULFID   107   126
FT   DISULFID   142   145
FT   DISULFID   206   218
FT   DISULFID   229   248
FT   DISULFID   264   267
FT   DISULFID   328   340
FT   DISULFID   351   370
FT   CARBOHYD    48    48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294   294       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   401   401       O-LINKED.
FT   CARBOHYD   533   533       O-LINKED.
FT   CARBOHYD   542   542       O-LINKED.
FT   CARBOHYD   546   546       O-LINKED.
FT   CARBOHYD   557   557       O-LINKED.
FT   CARBOHYD   571   571       O-LINKED.
FT   CARBOHYD   577   577       O-LINKED.
FT   CARBOHYD   593   593       O-LINKED.
FT   CARBOHYD   628   628       O-LINKED.
FT   VARSPLIC   402   427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                              PPKAGAEPASEREVS (in isoform LMW).
FT                              /FTId=VSP_001261.
FT   VARSPLIC   428   644       Missing (in isoform LMW).
FT                              /FTId=VSP_001262.
FT   CONFLICT   593   593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;

  Query Match             90.0%;  Score 618;  DB 1;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-51;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368


RESULT 2
KNL1_BOVIN
ID   KNL1_BOVIN     STANDARD;      PRT;   436 AA.
AC   P01046;
DT   21-JUL-1986 (Rel. 01, Created)
```

```
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen, LMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (3) the
CC       active peptide kallidin that is released from LMW-kininogen shows
CC       a variety of physiological effects: (3A) influence in smooth
CC       muscle contraction, (3B) induction of hypotension, (3C)
CC       natriuresis and diuresis (kidney).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=LMW I;
CC         IsoId=P01046-1; Sequence=Displayed;
CC       Name=HMW I;
CC         IsoId=P01044-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: LMW-kininogen is in contrast to HMW-kininogen not
CC       involved in blood clotting.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00426; CAA23709.1; -.
DR   PIR; A01283; KGBOL1.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     18
FT   CHAIN        19    436       KININOGEN, LMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    436       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    406       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
FT   CONFLICT    295    295       A -> T (IN REF. 1; CAA23709).
SQ   SEQUENCE   436 AA;  48427 MW;  F01F7EB6814BCE6C CRC64;

  Query Match             64.0%;  Score 440;  DB 1;  Length 436;
  Best Local Similarity   70.4%;  Pred. No. 1.8e-34;
  Matches   81;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              |||||||||::| ||||: || :||:||| |:|:| |||||:: |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         64 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              | || | |:||||||| |||||:||| |  | | |:| ||||||||:|:||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 367


RESULT 3
KNH1_BOVIN
ID   KNH1_BOVIN     STANDARD;      PRT;   621 AA.
AC   P01044;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 378-393.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 458-498.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
```

```
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW I;
CC         IsoId=P01044-1; Sequence=Displayed;
CC       Name=LMW I;
CC         IsoId=P01046-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01491; CAA24735.1; -.
DR   PIR; A01281; KGBOH1.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation;
KW   Inflammatory response; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     18       PROBABLE.
FT   CHAIN        19    621       KININOGEN, HMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    621       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    591       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
SQ   SEQUENCE   621 AA;  68890 MW;  D16850BEFE3C55CD CRC64;

  Query Match             64.0%;  Score 440;  DB 1;  Length 621;
  Best Local Similarity   70.4%;  Pred. No. 2.7e-34;
  Matches   81;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              |||||||::| |||: || :||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         64 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 367


RESULT 4
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (3) the
CC       active peptide kallidin that is released from LMW-kininogen shows
CC       a variety of physiological effects: (3A) influence in smooth
CC       muscle contraction, (3B) induction of hypotension, (3C)
CC       natriuresis and diuresis (kidney).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=LMW II;
CC         IsoId=P01047-1; Sequence=Displayed;
CC       Name=HMW II;
CC         IsoId=P01045-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: LMW-kininogen is in contrast to HMW-kininogen not
CC       involved in blood clotting.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
```

```
KW   Thiol protease inhibitor; Bradykinin; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1    18
FT   CHAIN        19   434       KININOGEN, LMW II.
FT   CHAIN        19   376       HEAVY CHAIN.
FT   PEPTIDE     378   386       BRADYKININ.
FT   CHAIN       387   434       LIGHT CHAIN.
FT   DOMAIN       19   135       CYSTATIN-LIKE 1.
FT   DOMAIN      136   256       CYSTATIN-LIKE 2.
FT   DOMAIN      257   376       CYSTATIN-LIKE 3.
FT   MOD_RES      19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87    87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136   136       O-LINKED (PARTIAL).
FT   CARBOHYD    168   168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197   197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204   204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280   280       N-LINKED (GLCNAC. . .).
FT   DISULFID     27   404       INTERCHAIN.
FT   DISULFID     82    93
FT   DISULFID    106   125
FT   DISULFID    141   144
FT   DISULFID    205   217
FT   DISULFID    228   247
FT   DISULFID    261   264
FT   DISULFID    325   337
FT   DISULFID    348   367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;

  Query Match             60.1%;  Score 413;  DB 1;  Length 434;
  Best Local Similarity   67.2%;  Pred. No. 6.6e-32;
  Matches   78;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              |:||:  |  :|||||: || :||:||| | |:| |||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 365


RESULT 5
KNH2_BOVIN
ID   KNH2_BOVIN     STANDARD;      PRT;   619 AA.
AC   P01045;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 376-391.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 387-455.
RX   MEDLINE=76260155; PubMed=956151;
RA   Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT   "Primary structure of bovine plasma high-molecular-weight kininogen.
RT   The amino acid sequence of a glycopeptide portion (fragment 1)
RT   following the C-terminus ot the bradykinin moiety.";
RL   J. Biochem. 79:1201-1222(1976).
RN   [5]
RP   SEQUENCE OF 456-496.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW II;
CC         IsoId=P01045-1; Sequence=Displayed;
CC       Name=LMW II;
CC         IsoId=P01047-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01492; CAA24736.1; -.
DR   PIR; A01282; KGBOH2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW   Inflammatory response; Pyrrolidone carboxylic acid.
FT   SIGNAL        1    18
FT   CHAIN        19   619       KININOGEN, HMW II.
FT   CHAIN        19   376       HEAVY CHAIN.
FT   PEPTIDE     378   386       BRADYKININ.
FT   CHAIN       387   619       LIGHT CHAIN.
FT   DOMAIN       19   135       CYSTATIN-LIKE 1.
FT   DOMAIN      136   256       CYSTATIN-LIKE 2.
```

```
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    400    400       O-LINKED.
FT   DISULFID     27    589       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
FT   VARIANT     398    398       T -> P.
FT   VARIANT     401    401       L -> V.
FT   VARIANT     454    454       H -> K.
SQ   SEQUENCE   619 AA;  68710 MW;  F04320A8EB0EE0DA CRC64;

  Query Match             60.1%;  Score 413;  DB 1;  Length 619;
  Best Local Similarity   67.2%;  Pred. No. 9.9e-32;
  Matches   78;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              |:||:   |  :|||||: || :||:||| | |:| |||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               | || | |:||||||| ||||||:|||    || | |:| |||||||:|:||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 365


RESULT 6
KNG_MOUSE
ID   KNG_MOUSE      STANDARD;      PRT;   661 AA.
AC   O08677; O08676; Q91XK5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX   MEDLINE=97342556; PubMed=9199253;
RA   Takano M., Kondo J., Yayama K., Otani M., Sano K., Okamoto H.;
RT   "Molecular cloning of cDNAs for mouse low-molecular-weight and high-
RT   molecular-weight prekininogens.";
RL   Biochim. Biophys. Acta 1352:222-230(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RC   STRAIN=C57BL/6J; TISSUE=Placenta;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
CC       clotting (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=O08677-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=O08677-2; Sequence=VSP_001263, VSP_001264;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D84435; BAA19743.1; -.
DR   EMBL; D84415; BAA19742.1; -.
DR   EMBL; AK005547; BAB24115.1; -.
DR   EMBL; BC018158; AAH18158.1; -.
DR   MGD; MGI:1097705; Kng.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    661       KININOGEN.
FT   CHAIN        19    379       KININOGEN HEAVY CHAIN.
FT   CHAIN       380    388       BRADYKININ.
FT   CHAIN       389    661       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    379       CYSTATIN-LIKE 3.
FT   DOMAIN      439    524       HIS-RICH.
FT   DISULFID     28    631       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    401    432       VSPPYIAREQEERDAETEQGPTHGHGWLHEKQ -> RLLRA
FT                                CEYKGRLSKAGAEPAPERQAESSQVKQ (in isoform
FT                                LMW).
FT                                /FTId=VSP_001263.
FT   VARSPLIC    433    661       Missing (in isoform LMW).
FT                                /FTId=VSP_001264.
SQ   SEQUENCE   661 AA;  73102 MW;  774460258D58796E CRC64;

  Query Match             60.1%;  Score 413;  DB 1;  Length 661;
  Best Local Similarity   66.7%;  Pred. No. 1.1e-31;
  Matches   78;  Conservative   11;  Mismatches   28;  Indels    0;  Gaps    0;

Qy          2 SGKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQV 61
              || | |:   | | ||||||| :||||:| | |:| :|||||:  ||:||| ||||  ||
Db        251 SGDDLVEALPKPCPGCPRDIPVDSPELKEVLGHSIAQLNAENDHPFYYKIDTVKKATSQV 310

Qy         62 VAGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||| || |:|:|||| |||||| || | || | |||||||| ||: ||| |: |||
Db        311 VAGTKYVIEFIARETKCSKESNTELAEDCEIKHLGQSLDCNANVYMRPWENKVVPTV 367


RESULT 7
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Buffalo;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
CC       clotting.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=P08934-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=P08934-2; Sequence=VSP_001265, VSP_001266;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW/LMW kininogens and two additional LMW-like
CC       kininogens: T-I and T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
```

```
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1     18
FT   CHAIN       19    639       KININOGEN.
FT   CHAIN       19    380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE     381   389       BRADYKININ.
FT   CHAIN       390   639       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19   136       CYSTATIN-LIKE 1.
FT   DOMAIN      137   258       CYSTATIN-LIKE 2.
FT   DOMAIN      259   380       CYSTATIN-LIKE 3.
FT   DOMAIN      439   514       HIS-RICH.
FT   DISULFID     28   609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83    94       BY SIMILARITY.
FT   DISULFID    107   126       BY SIMILARITY.
FT   DISULFID    142   145       BY SIMILARITY.
FT   DISULFID    206   218       BY SIMILARITY.
FT   DISULFID    229   248       BY SIMILARITY.
FT   DISULFID    264   267       BY SIMILARITY.
FT   DISULFID    328   340       BY SIMILARITY.
FT   DISULFID    351   370       BY SIMILARITY.
FT   CARBOHYD     82    82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127   127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294   294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    529   529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    402   433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                               CEYKGRLLKAGAGPAPERQAEASTVTP (in isoform
FT                               LMW).
FT                               /FTId=VSP_001265.
FT   VARSPLIC    434   639       Missing (in isoform LMW).
FT                               /FTId=VSP_001266.
FT   CONFLICT     61    61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;

  Query Match             59.7%;  Score 410;  DB 1;  Length 639;
  Best Local Similarity   66.4%;  Pred. No. 2e-31;
  Matches   77;  Conservative   13;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   : | ||||:|| :||||:| | |:| :|||||| ||||||| ||||  |||
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         63 AGKKYPIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              || ||  |:|:|||| |||||| |||  ||||:|||||:||| ||: ||| |: |||
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTV 368


RESULT 8
KNT2_RAT
ID   KNT2_RAT       STANDARD;      PRT;   430 AA.
AC   P08932;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
CC   -!- FUNCTION: Kininogens are plasma glycoproteins with a number of
CC       functions: (1) as precursor of the active peptide bradykinin they
CC       effect smooth muscle contraction, induction of hypotension and
CC       increase of vascular permeability. (2) They play a role in blood
CC       coagulation by helping to position optimally prekallikrein and
CC       factor XI next to factor XII. (3) They are inhibitor of thiol
CC       proteases.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- INDUCTION: In response to an inflammatory stimulant. T-kininogen
CC       II synthesis is induced and the plasma concentration of
CC       T-kininogen I is raised.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW and LMW kininogens produced by alternative splicing
CC       of the same gene, and two additional LMW-like kininogens: T-I and
CC       T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11885; AAA41491.1; -.
DR   PIR; B28055; B28055.
DR   GlycoSuiteDB; P08932; -.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL       1     18
FT   CHAIN       19    430       KININOGEN, T-II.
FT   CHAIN       19    375       HEAVY CHAIN.
FT   PEPTIDE     376   386       T-KININ.
FT   CHAIN       387   430       LIGHT CHAIN.
FT   DOMAIN       19   135       CYSTATIN-LIKE 1.
FT   DOMAIN      136   257       CYSTATIN-LIKE 2.
FT   DOMAIN      258   375       CYSTATIN-LIKE 3.
FT   DISULFID     28   404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83    94       BY SIMILARITY.
FT   DISULFID    107   125       BY SIMILARITY.
FT   DISULFID    141   144       BY SIMILARITY.
FT   DISULFID    205   217       BY SIMILARITY.
FT   DISULFID    228   247       BY SIMILARITY.
FT   DISULFID    263   266       BY SIMILARITY.
FT   DISULFID    327   339       BY SIMILARITY.
FT   DISULFID    350   369       BY SIMILARITY.
FT   CARBOHYD     82    82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126   126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168   168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204   204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326   326       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   430 AA;  47524 MW;  43EDF02D1BF55076 CRC64;

  Query Match             56.5%;  Score 388;  DB 1;  Length 430;
```

```
  Best Local Similarity   62.1%;  Pred. No. 1.6e-29;
  Matches   72;  Conservative   15;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |       | | |||::|| :||||:| | |:| :|||::|  ||||| ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              || || |:|:|||| |||::| |||  |||| |||||:||| ||: ||| |: |||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 9
KNT1_RAT
ID   KNT1_RAT       STANDARD;      PRT;   430 AA.
AC   P01048; P04081;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   T-kininogen I precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
GN   MAP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [2]
RP   SEQUENCE OF 5-430 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=86008266; PubMed=2413019;
RA   Anderson K.P., Heath E.C.;
RT   "The relationship between rat major acute phase protein and the
RT   kininogens.";
RL   J. Biol. Chem. 260:12065-12071(1985).
RN   [3]
RP   SEQUENCE OF 7-430 FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha 1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
RN   [4]
RP   SEQUENCE OF 1-65 FROM N.A.
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
CC   -!- FUNCTION: Kininogens are plasma glycoproteins with a number of
CC       functions: (1) as precursor of the active peptide bradykinin they
CC       effect smooth muscle contraction, induction of hypotension and
CC       increase of vascular permeability. (2) They play a role in blood
CC       coagulation by helping to position optimally prekallikrein and
CC       factor XI next to factor XII. (3) They are inhibitor of thiol
CC       proteases.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- INDUCTION: In response to an inflammatory stimulant. T-kininogen
CC       II synthesis is induced and the plasma concentration of
CC       T-kininogen I is raised.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW and LMW kininogens produced by alternative splicing
CC       of the same gene, and two additional LMW-like kininogens: T-I and
CC       T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   -!- CAUTION: In addition to the conflicts described in the feature
CC       table, Ref.2 sequence differs from that shown in positions 257,
CC       262, 268, 269, 295, 314, 315, 331, 332 and 389. In all those
CC       positions the alternate amino acid is the one present in T-II
CC       kininogen.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11883; AAA41489.1; -.
DR   EMBL; M11661; AAA41570.1; -.
DR   EMBL; M16454; AAA41568.1; -.
DR   EMBL; X02299; CAA26162.1; ALT_SEQ.
DR   PIR; A01286; KGRTT1.
DR   PIR; A23897; A23897.
DR   PIR; A27115; A27115.
DR   GlycoSuiteDB; P01048; -.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL       1     18
FT   CHAIN       19    430       KININOGEN, T-I.
FT   CHAIN       19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     26     28       LNC -> MDR (IN REF. 2).
FT   CONFLICT     55     55       V -> L (IN REF. 2).
FT   CONFLICT     61     61       E -> K (IN REF. 1).
FT   CONFLICT     83     83       C -> Y (IN REF. 3).
FT   CONFLICT    166    166       S -> F (IN REF. 2 AND 3).
FT   CONFLICT    179    181       REV -> TKI (IN REF. 2).
FT   CONFLICT    193    193       N -> D (IN REF. 2).
FT   CONFLICT    212    212       S -> F (IN REF. 2).
FT   CONFLICT    214    214       R -> H (IN REF. 3).
FT   CONFLICT    229    229       T -> R (IN REF. 2).
FT   CONFLICT    233    233       H -> Y (IN REF. 2).
FT   CONFLICT    257    257       E -> S (IN REF. 2).
FT   CONFLICT    262    262       N -> K (IN REF. 2).
FT   CONFLICT    264    264       R -> F (IN REF. 2).
FT   CONFLICT    268    269       RE -> KN (IN REF. 2).
FT   CONFLICT    295    295       I -> L (IN REF. 2).
FT   CONFLICT    314    315       VI -> TK (IN REF. 2).
FT   CONFLICT    331    332       SK -> TN (IN REF. 2).
FT   CONFLICT    389    389       R -> Q (IN REF. 2).
```

```
FT   CONFLICT    414    414       R -> G (IN REF. 2 AND 3).
FT   CONFLICT    415    415       A -> L (IN REF. 2).
FT   CONFLICT    420    421       DH -> ER (IN REF. 3).
FT   CONFLICT    430    430       P -> S (IN REF. 1).
SQ   SEQUENCE   430 AA;  47715 MW;  FAEBB78FAF4723C3 CRC64;

  Query Match             55.3%;  Score 380;  DB 1;  Length 430;
  Best Local Similarity   62.1%;  Pred. No. 9e-29;
  Matches   72;  Conservative   14;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              |  |   :   | | ||||:|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||  | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: |||
Db        312 AGVIYYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 10
CYTF_MOUSE
ID   CYTF_MOUSE     STANDARD;      PRT;   144 AA.
AC   O89098;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin F precursor (Leukocystatin) (Cystatin 7) (Cystatin-like
DE   metastasis-associated protein) (CMAP).
GN   CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98298157; PubMed=9632704;
RA   Halfon S., Ford J., Foster J., Dowling L., Lucian L., Sterling M.,
RA   Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,
RA   Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.;
RT   "Leukocystatin, a new class II cystatin expressed selectively by
RT   hematopoietic cells.";
RL   J. Biol. Chem. 273:16400-16408(1998).
CC   -!- FUNCTION: Inhibits papain and cathepsin L but with affinities
CC       lower than other cystatins. May play a role in immune regulation
CC       through inhibition of a unique target in the hematopoietic system.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF031826; AAC40140.1; -.
DR   EMBL; AF031825; AAC40139.1; -.
DR   HSSP; P01034; 1G96.
DR   MGD; MGI:1298217; Cst7.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW   Thiol protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    144       CYSTATIN F.
FT   ACT_SITE     36     36       REACTIVE SITE.
FT   SITE         80     84       SECONDARY AREA OF CONTACT.
FT   DISULFID     98    109       BY SIMILARITY.
FT   DISULFID    123    143       BY SIMILARITY.
SQ   SEQUENCE   144 AA;  16380 MW;  B5837334C1B4A89C CRC64;

  Query Match             25.0%;  Score 171.5;  DB 1;  Length 144;
  Best Local Similarity   35.5%;  Pred. No. 1.7e-09;
  Matches   39;  Conservative   22;  Mismatches   42;  Indels    7;  Gaps    3;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              || :   : :  | |: | ||:| : :   |:: | |    |  | || :| || ||||
Db         27 KDLI---SSVKPGFPKTIETNNPGVLKAARHSVEKFNNCTNDIFLFKESHVSKALVQVVK 83

Qy         64 GKKYFIDFVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPW 110
              | || ::     ||| |  : :| ::|:      | ::| | :||:|:||
Db         84 GLKYMLEVKIGRTTCRKTMHHQL-DNCDFQTNPALKRTLYCYSEVWVIPW 132


RESULT 11
CYTF_HUMAN
ID   CYTF_HUMAN     STANDARD;      PRT;   145 AA.
AC   O76096; Q9UED4;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cystatin F precursor (Leukocystatin) (Cystatin 7) (Cystatin-like
DE   metastasis-associated protein) (CMAP).
GN   CST7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98406133; PubMed=9733783;
RA   Ni J., Fernandez M.A., Danielsson L., Chillakuru R.A., Zhang J.,
RA   Grubb A., Su J., Gentz R., Abrahamson M.;
RT   "Cystatin F is a glycosylated human low molecular weight cysteine
RT   proteinase inhibitor.";
RL   J. Biol. Chem. 273:24797-24804(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98298157; PubMed=9632704;
RA   Halfon S., Ford J., Foster J., Dowling L., Lucian L., Sterling M.,
RA   Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,
RA   Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.;
RT   "Leukocystatin, a new class II cystatin expressed selectively by
RT   hematopoietic cells.";
RL   J. Biol. Chem. 273:16400-16408(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Morita M., Arakawa H., Yoshiuchi N.;
RT   "Human homologue of murine CMAP.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20399571; PubMed=10945474;
RA   Morita M., Hara Y., Tamai Y., Arakawa H., Nishimura S.;
RT   "Genomic construct and mapping of the gene for CMAP
RT   (Leukocystatin/Cystatin F, CST7) and identification of a proximal
RT   novel gene, BSCv (C20orf3).";
RL   Genomics 67:87-91(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
```

```
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Inhibits papain and cathepsin L but with affinities
CC       lower than other cystatins. May play a role in immune regulation
CC       through inhibition of a unique target in the hematopoietic system.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- TISSUE SPECIFICITY: Primarily expressed in peripheral blood cells
CC       and spleen.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF036342; AAC35747.1; -.
DR   EMBL; AF031824; AAC39788.1; -.
DR   EMBL; AB015225; BAA34941.1; ALT_INIT.
DR   EMBL; AB029636; BAB11886.1; ALT_INIT.
DR   EMBL; AL035661; CAB75498.1; -.
DR   EMBL; BC015507; AAH15507.1; ALT_INIT.
DR   HSSP; P01034; 1G96.
DR   Genew; HGNC:2479; CST7.
DR   MIM; 603253; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW   Thiol protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL       1     19       POTENTIAL.
FT   CHAIN       20    145       CYSTATIN F.
FT   ACT_SITE    37     37       REACTIVE SITE.
FT   SITE        81     85       SECONDARY AREA OF CONTACT.
FT   DISULFID    99    110       BY SIMILARITY.
FT   DISULFID   124    144       BY SIMILARITY.
FT   CARBOHYD    62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   145 AA;  16454 MW;  B2BCC4F76857CB0F CRC64;

  Query Match             23.8%;  Score 163.5;  DB 1;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.9e-09;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |  :  | : |:       | |:| | :||:|||         : |:|
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 12
CYTC_MACMU
ID   CYTC_MACMU     STANDARD;      PRT;   146 AA.
AC   O19092;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin C precursor.
GN   CST3.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97054523; PubMed=8898820;
RA   Wei L.H., Walker L.C., Levy E.;
RT   "Cystatin C. Icelandic-like mutation in an animal model of
RT   cerebrovascular beta-amyloidosis.";
RL   Stroke 27:2080-2085(1996).
CC   -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC       thought to serve an important physiological role as a local
CC       regulator of this enzyme activity.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U51912; AAB64050.1; -.
DR   HSSP; P01034; 1G96.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Amyloid; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    146       CYSTATIN C.
FT   ACT_SITE    37     37       REACTIVE SITE.
FT   SITE        81     85       SECONDARY AREA OF CONTACT.
FT   DISULFID    99    109       BY SIMILARITY.
FT   DISULFID   123    143       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  15857 MW;  F0B3BB774A29DF26 CRC64;

  Query Match             20.2%;  Score 138.5;  DB 1;  Length 146;
  Best Local Similarity   27.9%;  Pred. No. 2.4e-06;
  Matches   34;  Conservative   25;  Mismatches   52;  Indels   11;  Gaps    4;
```

```
Qy          8 QPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKY 67
              :||     || | |        :   |   ::: |  :|  :: :   | :|| |:||| |
Db         31 KPPR--LVGGPMDASVEEEGVRRALDFAVSEYNKASNDMYHSRALQVVRARKQIVAGVNY 88

Qy         68 FIDFVARETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
              |:|      |||:|   :   ::|   |    | :   |: ::| |||:      |:|::
Db         89 FLDVELGRTTCTK--TQPNLDNCPFHEQPHLKRKAFCSFQIYTVPWQ----GTMTLSKST 142

Qy        125 CE 126
              |:
Db        143 CQ 144


RESULT 13
CYTM_HUMAN
ID   CYTM_HUMAN     STANDARD;      PRT;   149 AA.
AC   Q15828;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Cystatin M precursor (Cystatin E).
GN   CST6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97150844; PubMed=8995380;
RA   Sotiropoulou G., Anisowicz A., Sager R.;
RT   "Identification, cloning, and characterization of cystatin M, a novel
RT   cysteine proteinase inhibitor, down-regulated in breast cancer.";
RL   J. Biol. Chem. 272:903-910(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97256812; PubMed=9099741;
RA   Ni J., Abrahamson M., Zhang M., Fernandez M.A., Grubb A., Su J.,
RA   Yu G.L., Li Y., Parmelee D., Xing L., Coleman T.A., Gentz S.,
RA   Thotakura R., Nguyen N., Hesselberg M., Gentz R.;
RT   "Cystatin E is a novel human cysteine proteinase inhibitor with
RT   structural resemblance to family 2 cystatins.";
RL   J. Biol. Chem. 272:10853-10858(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX   MEDLINE=21246880; PubMed=11348457;
RA   Zeeuwen P.L., Van Vlijmen-Willems I.M., Jansen B.J., Sotiropoulou G.,
RA   Curfs J.H., Meis J.F., Janssen J.J., Van Ruissen F., Schalkwijk J.;
RT   "Cystatin M/E expression is restricted to differentiated epidermal
RT   keratinocytes and sweat glands: a new skin-specific proteinase
RT   inhibitor that is a target for cross-linking by transglutaminase.";
RL   J. Invest. Dermatol. 116:693-701(2001).
CC   -!- FUNCTION: Shows moderate inhibition of cathepsin B but is not
CC       active against cathepsin C.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Restricted to the stratum granulosum of normal
CC       skin, the stratum granulosum/spinosum of psoriatic skin, and the
CC       secretory coils of eccrine sweat glands. Low expression levels are
CC       found in the nasal cavity.
CC   -!- PTM: Substrate for transglutaminases. Acts as an acyl acceptor but
CC       not as an acyl donor.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U62800; AAB06566.1; -.
DR   EMBL; U81233; AAB61305.1; -.
DR   EMBL; BC031334; AAH31334.1; -.
DR   HSSP; P01038; 1CEW.
DR   Genew; HGNC:2478; CST6.
DR   MIM; 601891; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Signal; Glycoprotein.
FT   SIGNAL       1     28       PROBABLE.
FT   CHAIN       29    149       CYSTATIN M.
FT   ACT_SITE    36     36       REACTIVE SITE.
FT   SITE        80     84       SECONDARY AREA OF CONTACT.
FT   DISULFID    98    113       BY SIMILARITY.
FT   DISULFID   126    146       BY SIMILARITY.
FT   CARBOHYD   137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   149 AA;  16511 MW;  2076A78BFC9FAC8C CRC64;

  Query Match             20.2%;  Score 138.5;  DB 1;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 2.4e-06;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              |  :   ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db         30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 14
CYTC_BOVIN
ID   CYTC_BOVIN     STANDARD;      PRT;   148 AA.
AC   P01035;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin C precursor (Colostrum thiol proteinase inhibitor).
GN   CST3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
RC   TISSUE=Cerebrospinal fluid, and Choroid plexus;
```

```
RX   MEDLINE=98094199; PubMed=9434110;
RA   Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bjoerk I.;
RT   "Molecular cloning and N-terminal analysis of bovine cystatin C
RT   identification of a full-length N-terminal region.";
RL   Biochim. Biophys. Acta 1343:203-210(1997).
RN   [2]
RP   SEQUENCE OF 37-148.
RX   MEDLINE=85231205; PubMed=3891407;
RA   Hirado M., Tsunasawa S., Sakiyama F., Niinobe M., Fujii S.;
RT   "Complete amino acid sequence of bovine colostrum low-Mr cysteine
RT   proteinase inhibitor.";
RL   FEBS Lett. 186:41-45(1985).
CC   -!- FUNCTION: This is a thiol proteinase inhibitor.
CC   -!- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y10811; CAA71771.1; -.
DR   HSSP; P01034; 1G96.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     30       PROBABLE.
FT   CHAIN        31    148       CYSTATIN C.
FT   MOD_RES      31     31       PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT   ACT_SITE     40     40       REACTIVE SITE.
FT   SITE         84     88       SECONDARY AREA OF CONTACT.
FT   DISULFID    102    112       BY SIMILARITY.
FT   DISULFID    126    146       BY SIMILARITY.
SQ   SEQUENCE   148 AA;  16265 MW;  EE740FE37CFB9F0E CRC64;

  Query Match             20.0%;  Score 137.5;  DB 1;  Length 148;
  Best Local Similarity   28.8%;  Pred. No. 3e-06;
  Matches   32;  Conservative   25;  Mismatches   35;  Indels   19;  Gaps    4;

Qy         24 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 83
              |   ::| |:  ::: |  :|  :  ::  | :|| |||:|  ||:|     |||:|  :
Db         48 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 105

Qy         84 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWECE 126
              :  :||        : :||     |: :||||||      |: :  : |:
Db        106 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 147


RESULT 15
FETB_RAT
ID   FETB_RAT       STANDARD;      PRT;   378 AA.
AC   Q9QX79;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Fetuin-B precursor (IRL685).
GN   FETUB.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=20407138; PubMed=10947975;
RA   Olivier E., Soury E., Ruminy P., Husson A., Parmentier F., Daveau M.,
RA   Salier J.-P.;
RT   "Fetuin-B, a second member of the fetuin family in mammals.";
RL   Biochem. J. 350:589-597(2000).
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- TISSUE SPECIFICITY: Liver.
CC   -!- SIMILARITY: Belongs to the fetuin family.
CC   -!- SIMILARITY: Contains 2 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ242926; CAB62543.1; -.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR001363; Fetuin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
KW   Glycoprotein; Signal; Repeat.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    378       FETUIN-B.
FT   DOMAIN       27    152       CYSTATIN-LIKE 1.
FT   DOMAIN      153    273       CYSTATIN-LIKE 2.
FT   DISULFID     96    107       BY SIMILARITY.
FT   DISULFID    120    140       BY SIMILARITY.
FT   DISULFID    154    157       BY SIMILARITY.
FT   DISULFID    217    224       BY SIMILARITY.
FT   DISULFID    237    260       BY SIMILARITY.
FT   CARBOHYD     40     40       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    139    139       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   378 AA;  41532 MW;  066C0A5C3B03C878 CRC64;

  Query Match             19.7%;  Score 135;  DB 1;  Length 378;
  Best Local Similarity   25.8%;  Pred. No. 1.5e-05;
  Matches   33;  Conservative   32;  Mismatches   49;  Indels   14;  Gaps    5;

Qy          7 VQPPTK-----ICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQV 61
              ::| :|     :|  ||  :  ::| : |  | :: | |:|| :  |  :  | ||  |
Db        142 LRPVSKRKIHSMCPDCPHPVDLSAPSVLEAATESLAKFNSENPSKQYALV-KVTKATTQW 200

Qy         62 VAGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPW--EKKIYPTVT 119
              | |  ||:::: :|: |::  :    :: ::: :|    |  :  |    ::   |||
Db        201 VVGPSYFVEYLIKESPCTQSQDSCSLQASDSEPVGL---CQGSLIKSPGVPPQRFKKTVT 257

Qy        120 VNHWECEF 127
              |:   |||
Db        258 VS---CEF 262


Search completed: September 24, 2004, 14:09:13
Job time : 9.636 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:06:08 ; Search time 13.716 Seconds
(without alignments)
890.662 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	618	90.0	427	1	KGHUL1	kininogen, LMW pre
2	618	90.0	644	1	KGHUH1	kininogen, HMW pre
3	440	64.0	436	1	KGBOL1	kininogen, LMW I p
4	440	64.0	621	1	KGBOH1	kininogen, HMW I p
5	413	60.1	434	1	KGBOL2	kininogen, LMW II
6	413	60.1	619	1	KGBOH2	kininogen, HMW II
7	410	59.7	433	2	A28055	K-kininogen, LMW I
8	410	59.7	639	2	A25486	kininogen, HMW I p
9	388	56.5	430	2	A23897	major acute phase
10	388	56.5	430	2	B28055	T-kininogen, LMW I
11	381	55.5	423	1	KGRTM	major acute phase
12	380	55.3	430	1	KGRTT1	T-kininogen I prec
13	137.5	20.0	112	1	UDBO	cystatin - bovine
14	132.5	19.3	146	1	UDHU	cystatin C precurs
15	130	18.9	127	2	S07085	cystatin C precurs
16	129	18.8	120	2	S10587	cystatin C - rat
17	128	18.6	111	2	A28793	cystatin - puff ad
18	127.5	18.6	140	2	A36163	cystatin C precurs
19	124.5	18.1	141	2	B29632	cystatin SA precur
20	118.5	17.2	139	1	UDCH	cystatin precursor
21	113	16.4	141	2	JQ1470	cystatin S precurs
22	112	16.3	91	2	S68034	T-kininogen (clone
23	112	16.3	91	2	S68035	T-kininogen (clone
24	111	16.2	111	1	JC2040	cystatin - chum sa
25	109.5	15.9	141	1	UDHUP2	cystatin SN precur
26	108.5	15.8	141	1	UDHUP1	cystatin S precurs
27	107	15.6	139	2	T33740	hypothetical prote
28	106	15.4	132	2	JC4918	cystatin precursor
29	105.5	15.4	162	2	A43428	onchocystatin - ne
30	102	14.8	122	2	A43644	sarcocystatin A pr
31	101.5	14.8	133	2	JC4536	cystatin precursor
32	97.5	14.2	135	2	JC4007	cystatin II - maiz
33	97	14.1	142	2	A47142	cystatin D precurs
34	95.5	13.9	143	2	T33301	hypothetical prote
35	94.5	13.8	345	2	S21094	alpha-2-HS-glycopr
36	93.5	13.6	139	2	A45361	cystatin-related e
37	92	13.4	100	2	JH0269	cystatin - avocado
38	91.5	13.3	135	2	S27239	cysteine proteinas
39	91.5	13.3	375	2	A32827	fetuin precursor -
40	88.5	12.9	125	2	T00752	cysteine proteinas
41	88.5	12.9	134	2	S54828	cysteine proteinas
42	88.5	12.9	134	2	JC4882	cystatin - maize
43	88	12.8	199	2	S65071	cystatin - field m
44	88	12.8	367	1	WOHU	alpha-2-HS-glycopr
45	84.5	12.3	592	2	T50516	Ca2+ dependent lip

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621; PMID:6441591
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582; PMID:2989293
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, New
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen.
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021; PMID:3365237
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic flui
A;Reference number: A31905; MUID:89034061; PMID:3182782
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
A;Reference number: A34030; MUID:88106632; PMID:3337729
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583; PMID:2989294
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622; PMID:4952632
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 90.0%; Score 618; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 3.6e-50;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368
```

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621; PMID:6441591
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582; PMID:2989293
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854; PMID:8467916
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininoge
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined in
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen. The
A;Reference number: A24871; MUID:86108361; PMID:3484703
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New York
A;Title: Amino acid sequence of the light chain of human high molecular mass kininogen.
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021; PMID:3365237
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
A;Reference number: A31905; MUID:89034061; PMID:3182782
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasm
A;Reference number: A34030; MUID:88106632; PMID:3337729
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
A;Reference number: S02482; MUID:89076517; PMID:3264507
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
A;Reference number: A61495; MUID:88211869; PMID:3366244
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>

```
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495
A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133; PMID:2013314
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593; PMID:7733867
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974; PMID:7589467
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583; PMID:2989294
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622; PMID:4952632
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

  Query Match             90.0%;  Score 618;  DB 1;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-50;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368


RESULT 3
KGBOL1
kininogen, LMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01283
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors and
A;Reference number: A93984; MUID:83117859; PMID:6572010
A;Accession: A01283
A;Molecule type: mRNA
A;Residues: 1-436 <NAW>
A;Cross-references: GB:J00010; GB:V00426; NID:g163256; PIDN:AAA30604.1; PID:g163257
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-436/Product: LMW kininogen I #status predicted <MAT>
F;19-378/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status predicted <BDY>
F;389-436/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27-406,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;47,87,168,169,197,204/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;378-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status predicted

  Query Match             64.0%;  Score 440;  DB 1;  Length 436;
  Best Local Similarity   70.4%;  Pred. No. 1.7e-33;
  Matches   81;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              |||||||::| |||: || :||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         64 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              | || | |:||||||| ||||||:||| || | |:| |||||||:|:|||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 367


RESULT 4
KGBOH1
kininogen, HMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01281; A91923; A91938; A29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininoge
A;Reference number: A93317; MUID:84014106; PMID:6571699
```

```
A;Accession: A01281
A;Molecule type: mRNA
A;Residues: 1-621 <KIT>
A;Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420; PMID:4986212
A;Accession: A91923
A;Molecule type: protein
A;Residues: 378-393 <KAT>
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265; PMID:1169237
A;Accession: A91938
A;Molecule type: protein
A;Residues: 458-498 <HAN>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530; PMID:3546295
A;Accession: A29559
A;Molecule type: protein
A;Residues: 'Z',20-123,'I',125-127,'I',129-378 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-621/Product: HMW prokininogen I #status predicted <MAT>
F;19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F;389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F;417-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;398,406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status ex
F;498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

  Query Match             64.0%;  Score 440;  DB 1;  Length 621;
  Best Local Similarity   70.4%;  Pred. No. 2.4e-33;
  Matches   81;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          4 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 63
              ||||||||::| |||: || :||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         64 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
                | || | |:|||||||| |||||||:|||     || | |:| |||||||:|:|||| |
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 367


RESULT 5
KGBOL2
kininogen, LMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C;Accession: A01284
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors and
A;Reference number: A93984; MUID:83117859; PMID:6572010
A;Accession: A01284
A;Molecule type: mRNA
A;Residues: 1-434 <NAW>
A;Cross-references: GB:V00427; GB:J00011; NID:g489; PIDN:CAA23710.1; PID:g490
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, ir
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyco
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-434/Product: LMW kininogen II #status predicted <MAT>
F;19-377/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status predicted <BDY>
F;387-434/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27-404,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bonds: #
F;47,87,168,169,197,204,280/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;376-377/Cleavage site: Met-Lys (kallikrein) #status predicted
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status predicted

  Query Match             60.1%;  Score 413;  DB 1;  Length 434;
  Best Local Similarity   67.2%;  Pred. No. 5.5e-31;
  Matches   78;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              |:||:  |  :|||||: || :||:||| | |:| |||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
                | || | |:|||||||| |||||||:|||     || | |:| |||||||:|:|||| |
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 365


RESULT 6
KGBOH2
kininogen, HMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01282; A91923; A91941; A91938; B29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininoge
A;Reference number: A93317; MUID:84014106; PMID:6571699
A;Accession: A01282
A;Molecule type: mRNA
A;Residues: 1-619 <KIT>
A;Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R;Kato, H.; Nagasawa, S.; Suzuki, T.
```

J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420; PMID:4986212
A;Accession: A91923
A;Molecule type: protein
A;Residues: 376-391 <KAT>
R;Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A;Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A;Reference number: A91941; MUID:76260155; PMID:956151
A;Accession: A91941
A;Molecule type: protein
A;Residues: 387-455 <HAN>
A;Note: 398-Pro, 401-Val, and 455-Lys were also found
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265; PMID:1169237
A;Accession: A91938
A;Molecule type: protein
A;Residues: 456-496 <HA2>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530; PMID:3546295
A;Accession: B29559
A;Molecule type: protein
A;Residues: 'Z',20-104,'E',106-256,'XX',257-376 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-619/Product: HMW kininogen II #status predicted <MAT>
F;19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F;387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F;418-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-589,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bonds:
F;47/Binding site: carbohydrate (Asn) (covalent) #status absent
F;87,168,169,204,280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;396,400,404,510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;397,398,518,522,534,546,551,568/Binding site: carbohydrate (Thr) (covalent) #status ex
F;496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 60.1%; Score 413; DB 1; Length 619;
Best Local Similarity 67.2%; Pred. No. 8.2e-31;
Matches 78; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

```
Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              |:||:  |  :|||||: || :||:||| | |:| ||||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:|||| 
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTV 365


RESULT 7
A28055
K-kininogen, LMW I precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C;Accession: A28055
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: A28055
A;Molecule type: mRNA
A;Residues: 1-433 <FUR>
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-433/Product: K-kininogen, LMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 59.7%; Score 410; DB 2; Length 433;
Best Local Similarity 66.4%; Pred. No. 1.1e-30;
Matches 77; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   : | ||||:|| :||||:| | |:| :|||||| ||||||| ||||  |||
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               || || |:|:|||| |||||| ||| ||||:|||||:||| ||: ||| |: |||
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTV 368


RESULT 8
A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-639/Product: kininogen, HMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 59.7%; Score 410; DB 2; Length 639;
Best Local Similarity 66.4%; Pred. No. 1.6e-30;
Matches 77; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   : | ||||:|| :||||:| | |:| :|||||| ||||||| ||||  |||
```

```
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              || || |:|:|||| |||||| ||| ||||:|||||:||| ||: ||| |: |||
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTV 368


RESULT 9
A23897
major acute phase alpha-1 protein (version 2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 20-Aug-1999
C;Accession: A23897; B23897
R;Anderson, K.P.; Heath, E.C.
J. Biol. Chem. 260, 12065-12071, 1985
A;Title: The relationship between rat major acute phase protein and the kininogens.
A;Reference number: A23897; MUID:86008266; PMID:2413019
A;Accession: A23897
A;Molecule type: protein
A;Residues: 1-14 <AND1>
A;Accession: B23897
A;Molecule type: mRNA
A;Residues: 5-430 <AND2>
A;Cross-references: GB:M11661; NID:g205307; PIDN:AAA41570.1; PID:g205308
A;Note: the authors translated the codon CTC for residue 410 as Arg, CTA for residue 415
C;Superfamily: kininogen; cystatin homology
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>

  Query Match           56.5%;  Score 388;  DB 2;  Length 430;
  Best Local Similarity  62.1%;  Pred. No. 1.2e-28;
  Matches   72;  Conservative   15;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |      | | |||::|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              || || |:|:|||| |||::|  |||  ||||  |||||:||| ||: ||| |: |||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 10
B28055
T-kininogen, LMW II precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Dec-1997
C;Accession: B28055; E25486; B28526; C28526
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: B28055
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: E25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226; PMID:3121623
A;Accession: B28526
A;Molecule type: protein
A;Residues: 'E',20-25,'MD',28-48,376-430 <ENJ>
A;Accession: C28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-388,'R',390-419,'ER',422-430 <EN2>
C;Superfamily: kininogen; cystatin homology
C;Keywords: glycoprotein; pyroglutamic acid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen, LMW II #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #status

  Query Match           56.5%;  Score 388;  DB 2;  Length 430;
  Best Local Similarity  62.1%;  Pred. No. 1.2e-28;
  Matches   72;  Conservative   15;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |      | | |||::|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              || || |:|:|||| |||::|  |||  ||||  |||||:||| ||: ||| |: |||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 11
KGRTM
major acute phase alpha-1 protein precursor - rat (fragment)
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 12-Apr-1996
C;Accession: A01285
R;Cole, T.; Inglis, A.S.; Roxburgh, C.M.; Howlett, G.J.; Schreiber, G.
FEBS Lett. 182, 57-61, 1985
A;Title: Major acute phase alpha1-protein of the rat is homologous to bovine kininogen ar
A;Reference number: A01285; MUID:85127561; PMID:2578992
A;Accession: A01285
A;Molecule type: mRNA
A;Residues: 1-423 <COL>
C;Comment: This plasma glycoprotein inhibits cysteine proteinases. During acute inflammat
r.
C;Superfamily: kininogen; cystatin homology
C;Keywords: bradykinin; cysteine proteinase inhibitor; duplication; glycoprotein; inflamm
F;1-11/Domain: signal sequence (fragment) #status predicted <SIG>
F;12-423/Product: major acute phase alpha-1 protein #status predicted <MAT>
F;12-123/Domain: cystatin homology <CY1>
F;134-245/Domain: cystatin homology <CY2>
F;256-367/Domain: cystatin homology <CY3>
F;371-379/Product: bradykinin #status predicted <BDY>
F;12/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;161,197/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match           55.5%;  Score 381;  DB 1;  Length 423;
  Best Local Similarity  62.1%;  Pred. No. 5.3e-28;
  Matches   72;  Conservative   14;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   | | |||:|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        245 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||  | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: |||
Db        305 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 360


RESULT 12
KGRTT1
T-kininogen I precursor - rat
N;Alternate names: 73K protein; LMW kininogen T-I
N;Contains: bradykinin; T-kinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 22-Jun-1999
C;Accession: A01286; D25486; A28526; PL0193; JQ0027; B25488; A28525; S68036
```

R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: A01286
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
A;Cross-references: GB:M11883; NID:g205084; PIDN:AAA41489.1; PID:g205085
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: D25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226; PMID:3121623
A;Accession: A28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-430 <ENJ>
R;Kanda, S.; Sugiyama, K.; Takahashi, M.; Shumiya, S.; Tomino, S.; Nagase, S.
Jpn. J. Cancer Res. 81, 63-68, 1990
A;Title: Identification of a protein increasing in serum of Nagase analbuminemic rats be
A;Reference number: PL0193; MUID:90216390; PMID:2108948
A;Accession: PL0193
A;Molecule type: mRNA
A;Residues: 330-420,'R',422-429,'P' <KAN>
R;Anderson, K.P.; Croyle, M.L.; Lingrel, J.B.
Gene 81, 119-128, 1989
A;Title: Primary structure of a gene encoding rat T-kininogen.
A;Reference number: JQ0027; MUID:90034172; PMID:2806908
A;Accession: JQ0027
A;Molecule type: DNA
A;Residues: 1-60,'E',62-113,'R',115-165,'F',167-178,'TKI',182-211,'F',213-256,'S',258-38
A;Experimental source: strain Sprague-Dawley
R;Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A;Title: Differing utilization of homologous transcription initiation sites of rat K and
A;Reference number: A25488; MUID:87137465; PMID:3818598
A;Accession: B25488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <KAG>
A;Cross-references: GB:M14356; NID:g205090; PIDN:AAA41492.1; PID:g205091
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 965-972, 1988
A;Title: Purification and characterization of two kinds of low molecular weight kininoge
A;Reference number: A28525; MUID:88087225; PMID:3335530
A;Accession: A28525
A;Molecule type: protein
A;Residues: 376-430 <EN2>
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of agi
A;Reference number: S68034; MUID:96032652; PMID:7574705
A;Accession: S68036
A;Molecule type: mRNA
A;Residues: 340-430 <SIE>
A;Experimental source: clone pSG17
C;Comment: At least three types of LMW kininogen precursors are present in rat plasma, t
ceding bradykinin.
C;Comment: T-kininogens contain T-kinin (I-S-bradykinin), a novel kinin isolated after t
d of an Arg or Lys, it is probably not released from its precursor by either tissue or p
C;Comment: The T-kininogens are produced in response to an inflammatory stimulant.
C;Genetics:
A;Introns: 65/3; 102/3; 130/1; 187/3; 223/2; 252/1; 309/3; 345/3; 374/3; 398/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: acute phase; bradykinin; cysteine proteinase inhibitor; duplication; glycopr
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen I #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;378-386/Product: bradykinin #status predicted <BDY>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #status

Query Match 55.3%; Score 380; DB 1; Length 430;
Best Local Similarity 62.1%; Pred. No. 6.7e-28;
Matches 72; Conservative 14; Mismatches 30; Indels 0; Gaps 0;

Qy 3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
Db 252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy 63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
Db 312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367

RESULT 13
UDBO
cystatin - bovine
N;Alternate names: thiol proteinase inhibitor
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 06-Dec-1996
C;Accession: A01271
R;Hirado, M.; Tsunasawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.
FEBS Lett. 186, 41-45, 1985
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase inh
A;Reference number: A01271; MUID:85231205; PMID:3891407
A;Accession: A01271
A;Molecule type: protein
A;Residues: 1-112 <HIR>
C;Superfamily: cystatin; cystatin homology
C;Keywords: colostrum; cysteine proteinase inhibitor
F;2-112/Domain: cystatin homology <CYS>
F;48-52/Region: inhibitory #status predicted
F;66-76,90-110/Disulfide bonds: #status predicted

Query Match 20.0%; Score 137.5; DB 1; Length 112;
Best Local Similarity 28.8%; Pred. No. 7.5e-06;
Matches 32; Conservative 25; Mismatches 35; Indels 19; Gaps 4;

Qy 24 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 83
Db 12 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 69

Qy 84 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWECE 126
Db 70 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 111

RESULT 14
UDHU
cystatin C precursor [validated] - human
N;Alternate names: gamma-CSF; gamma-trace; neuroendocrine basic polypeptide; post-gamma
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 31-Mar-1991 #text_change 08-Dec-2000
C;Accession: S10216; S00004; JL0095; A33400; S02751; A01270; A25434; S12288; A32732; A60
R;Abrahamson, M.; Olafsson, I.; Palsdottir, A.; Ulvsbaeck, M.; Lundwall, A.; Jensson, O.
Biochem. J. 268, 287-294, 1990
A;Title: Structure and expression of the human cystatin C gene.
A;Reference number: S10216; MUID:90303202; PMID:2363674
A;Accession: S10216
A;Molecule type: DNA
A;Residues: 1-146 <AB1>
A;Cross-references: EMBL:X52255; NID:g30257; PIDN:CAA36497.1; PID:g296643
R;Abrahamson, M.; Grubb, A.; Olafsson, I.; Lundwall, A.
FEBS Lett. 216, 229-233, 1987
A;Title: Molecular cloning and sequence analysis of cDNA coding for the precursor of the
A;Reference number: S00004; MUID:87219149; PMID:3495457

A;Accession: S00004
A;Molecule type: mRNA
A;Residues: 1-146 <AB2>
A;Cross-references: EMBL:X05607; NID:g30371; PIDN:CAA29096.1; PID:g755738
R;Levy, E.; Lopez-Otin, C.; Ghiso, J.; Geltner, D.; Frangione, B.
J. Exp. Med. 169, 1771-1778, 1989
A;Title: Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a
A;Reference number: JL0095; MUID:89235594; PMID:2541223
A;Accession: JL0095
A;Molecule type: DNA
A;Residues: 1-146 <LEV>
A;Cross-references: GB:X61681; NID:g30367; PIDN:CAA43856.2; PID:g4490944
A;Note: the cystatin C gene isolated from the brain of an Icelandic patient with heredit
e)
R;Saitoh, E.; Sabatini, L.M.; Eddy, R.L.; Shows, T.B.; Azen, E.A.; Isemura, S.; Sanada,
Biochem. Biophys. Res. Commun. 162, 1324-1331, 1989
A;Title: The human cystatin C gene (CST3) is a member of the cystatin gene family which
A;Reference number: A33400; MUID:89350949; PMID:2764935
A;Accession: A33400
A;Molecule type: DNA
A;Residues: 1-24,'T',26-146 <SAI>
A;Cross-references: GB:M27889; GB:M27890; GB:M27891; NID:g181385; PIDN:AAA52164.1; PID:g
R;Ghiso, J.; Cowan, N.; Frangione, B.
Biol. Chem. Hoppe-Seyler 369, 205-208, 1988
A;Title: Isolation of a sequence encoding human cystatin C. Conservation of exon-intron
A;Reference number: S02751; MUID:89076507; PMID:3264504
A;Accession: S02751
A;Molecule type: DNA
A;Residues: 82-119 <GH2>
A;Cross-references: EMBL:M27769
A;Note: the authors translated the codon ACC for residue 105 as Thr; the sequence shown
R;Grubb, A.; Lofberg, H.
Proc. Natl. Acad. Sci. U.S.A. 79, 3024-3027, 1982
A;Title: Human gamma-trace, a basic microprotein: amino acid sequence and presence in th
A;Reference number: A01270; MUID:82222268; PMID:6283552
A;Accession: A01270
A;Molecule type: protein
A;Residues: 27-131,'S',133-146 <GRU>
R;Ghiso, J.; Jensson, O.; Frangione, B.
Proc. Natl. Acad. Sci. U.S.A. 83, 2974-2978, 1986
A;Title: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Iceland t
A;Reference number: A25434; MUID:86206076; PMID:3517880
A;Accession: A25434
A;Molecule type: protein
A;Residues: 37-93,'Q',95-146 <GHI>
R;Turk, V.; Brzin, J.; Longer, M.; Ritonja, A.; Eropkin, M.; Borchart, U.; Machleidt, W.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1487-1496, 1983
A;Title: Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystati
A;Reference number: S01461; MUID:84110059; PMID:6662498
A;Accession: S12288
A;Molecule type: protein
A;Residues: 27-73 <TUR>
R;Brzin, J.; Popovic, T.; Turk, V.
Biochem. Biophys. Res. Commun. 118, 103-109, 1984
A;Title: Human cystatin, a new protein inhibitor of cysteine proteinases.
A;Reference number: A32732; MUID:84128015; PMID:6365094
A;Accession: A32732
A;Molecule type: protein
A;Residues: 27-76 <BRZ>
R;Olafsson, I.; Gudmundsson, G.; Abrahamson, M.; Jensson, O.; Grubb, A.
Scand. J. Clin. Lab. Invest. 50, 85-93, 1990
A;Title: The amino terminal portion of cerebrospinal fluid cystatin C in hereditary cyst
A;Reference number: A60552; MUID:90193615; PMID:2315647
A;Accession: A60552
A;Molecule type: protein
A;Residues: 27-49,'XX',52-64 <OLA>
A;Note: this protein, purified from cerebrospinal fluid of patients with the autosomal d
e defective gene is not present in CSF but is found instead in amyloid deposits
R;Popovic, T.; Brzin, J.; Ritonja, A.; Turk, V.
Biol. Chem. Hoppe-Seyler 371, 575-580, 1990
A;Title: Different forms of human cystatin C.
A;Reference number: S10607; MUID:91025625; PMID:2222856
A;Accession: S10607
A;Molecule type: protein
A;Residues: 27-53 <POP>
A;Experimental source: urine, kidney disease
A;Note: truncated forms with amino ends at positions 35 and 36 of the precursor were also
R;Grubb, A.; Lofberg, H.; Barrett, A.J.
FEBS Lett. 170, 370-374, 1984
A;Title: The disulphide bridges of human cystatin C (gamma-trace) and chicken cystatin.
A;Reference number: S01462
A;Contents: annotation; disulfide bonds
R;Berti, P.J.; Storer, A.C.
Biochem. J. 302, 411-416, 1994
A;Title: Local pH-dependent conformational changes leading to proteolytic susceptibility
A;Reference number: S55305; MUID:94379969; PMID:8092991
A;Accession: S55305
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-49;106-146 <BER>
C;Comment: This protein is found in the post-gamma-globulin fraction of cerebrospinal flu
f patients with certain autoimmune diseases.
C;Comment: This protein is an inhibitor of cysteine proteinases and may serve an importar
C;Comment: A mutant cystatin C, with 94-Gln, is deposited in hereditary cerebral hermorrh
C;Genetics:
A;Gene: GDB:CST3
A;Cross-references: GDB:119817; OMIM:105150
A;Map position: 20p11.2-20p11.2
A;Introns: 81/3; 119/3
C;Superfamily: cystatin; cystatin homology
C;Keywords: amyloid; cysteine proteinase inhibitor; extracellular protein; hydroxyproline
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-146/Product: cystatin C #status experimental <MAT>
F;35-146/Domain: cystatin homology <CYS>
F;81-85/Region: inhibitory #status predicted
F;29/Modified site: hydroxyproline (Pro) (partial) #status experimental
F;99-109,123-143/Disulfide bonds: #status experimental

Query Match 19.3%; Score 132.5; DB 1; Length 146;
Best Local Similarity 27.0%; Pred. No. 2.9e-05;
Matches 33; Conservative 25; Mismatches 53; Indels 11; Gaps 4;

```
Qy          8 QPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKY 67
              :||    || | |       :   |   : : |  :|  :: :   | :|| |:|||  |
Db         31 KPPR--LVGGPMDASVEEEGVRRALDFAVGEYNKASNDMYHSRALQVVRARKQIVAGVNY 88

Qy         68 FIDFVARETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
              |:|     |||:|   :   ::|   :   | :   |: ::| |||:     |:|::
Db         89 FLDVELGRTTCTK--TQPNLDNCPFHDQPHLKRKAFCSFQIYAVPWQ----GTMTLSKST 142

Qy        125 CE 126
              |:
Db        143 CQ 144
```

RESULT 15
S07085
cystatin C precursor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 16-Jul-1999
C;Accession: S07085; S01337; S21109
R;Cole, T.; Dickson, P.W.; Esnard, F.; Averill, S.; Risbridger, G.P.; Gauthier, F.; Schre
Eur. J. Biochem. 186, 35-42, 1989
A;Title: The cDNA structure and expression analysis of the genes for the cysteine protein
A;Reference number: S07085; MUID:90092122; PMID:2689174
A;Accession: S07085
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-127 <COL>
A;Cross-references: EMBL:X16957; NID:g56041; PIDN:CAA34831.1; PID:g736290
R;Esnard, A.; Esnard, F.; Faucher, D.; Gauthier, F.
FEBS Lett. 236, 475-478, 1988
A;Title: Two rat homologues of human cystatin C.
A;Reference number: S01337; MUID:88313020; PMID:3044831
A;Accession: S01337
A;Molecule type: protein

A;Residues: 8-49 <ESN>
R;Esnard, A.; Esnard, F.; Guillou, F.; Gauthier, F.
FEBS Lett. 300, 131-135, 1992
A;Title: Production of the cysteine proteinase inhibitor cystatin C by rat Sertoli cells
A;Reference number: S21109; MUID:92225121; PMID:1563513
A;Accession: S21109
A;Molecule type: protein
A;Residues: 8,'XX',11-20 <ES2>
C;Superfamily: cystatin; cystatin homology
C;Keywords: cysteine proteinase inhibitor
F;16-127/Domain: cystatin homology <CYS>
F;80-90,104-124/Disulfide bonds: #status predicted

Query Match 18.9%; Score 130; DB 2; Length 127;
Best Local Similarity 28.0%; Pred. No. 4.3e-05;
Matches 30; Conservative 28; Mismatches 43; Indels 6; Gaps 4;

```
Qy          8 QPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKY 67
              :|| :: :| |:: :  :   ::  |  ::: |  :|  :: :   | :|| |:|||  |
Db         11 RPPPRL-LGAPQEADASEEGVQRALDFAVSEYNKGSNDAYHSRAIQVVRARKQLVAGINY 69

Qy         68 FIDFVARETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWE 111
              ::|   |||:| |   || :|   :   | :   |: ::| |||:
Db         70 YLDVEMGRTTCTK-SQTNLT-NCPFHDQPHLMRKALCSFQIYSVPWK 114
```

Search completed: September 24, 2004, 14:10:49
Job time : 14.716 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:08:41 ; Search time 44.704 Seconds
(without alignments)
913.519 Million cell updates/sec

Title: US-10-661-784-3
Perfect score: 687
Sequence: 1 GSGKDFVQPPTKICVGCPRD.....VPWEKKIYPTVTVNHWECEF 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	618	90.0	390	15	US-10-162-335-82	Sequence 82, Appl
2	618	90.0	398	15	US-10-162-335-70	Sequence 70, Appl
3	618	90.0	427	10	US-09-919-039-29	Sequence 29, Appl
4	618	90.0	615	15	US-10-162-335-72	Sequence 72, Appl
5	618	90.0	644	15	US-10-162-335-74	Sequence 74, Appl
6	618	90.0	644	15	US-10-162-335-84	Sequence 84, Appl
7	381	55.5	424	14	US-10-316-253-217	Sequence 217, App
8	380	55.3	430	14	US-10-316-253-215	Sequence 215, App
9	163.5	23.8	145	14	US-10-329-428-2	Sequence 2, Appli
10	163.5	23.8	167	10	US-09-746-783-197	Sequence 197, App
11	163	23.7	178	9	US-09-969-834-1	Sequence 1, Appli
12	138.5	20.2	121	9	US-09-775-932-14	Sequence 14, Appl
13	138.5	20.2	128	9	US-09-775-932-12	Sequence 12, Appl
14	138.5	20.2	149	9	US-09-940-497-2	Sequence 2, Appli
15	137.5	20.0	112	8	US-08-849-303-16	Sequence 16, Appl
16	137.5	20.0	112	16	US-10-655-136-16	Sequence 16, Appl
17	136.5	19.9	118	9	US-09-775-932-24	Sequence 24, Appl
18	132.5	19.3	120	9	US-09-775-932-2	Sequence 2, Appli
19	132.5	19.3	120	16	US-10-695-194-2	Sequence 2, Appli
20	132.5	19.3	146	8	US-08-849-303-17	Sequence 17, Appl
21	132.5	19.3	146	9	US-09-940-497-3	Sequence 3, Appli
22	132.5	19.3	146	9	US-09-969-834-3	Sequence 3, Appli
23	132.5	19.3	146	14	US-10-329-428-3	Sequence 3, Appli
24	132.5	19.3	146	14	US-10-376-564-47	Sequence 47, Appl
25	132.5	19.3	146	16	US-10-655-136-17	Sequence 17, Appl
26	132.5	19.3	146	16	US-10-695-194-1	Sequence 1, Appli
27	132.5	19.3	249	16	US-10-257-384A-4	Sequence 4, Appli
28	132.5	19.3	641	16	US-10-257-384A-2	Sequence 2, Appli
29	131.5	19.1	317	12	US-10-210-172-82	Sequence 82, Appl
30	131.5	19.1	345	12	US-10-210-172-86	Sequence 86, Appl
31	131.5	19.1	356	12	US-10-210-172-84	Sequence 84, Appl
32	131.5	19.1	369	12	US-10-210-172-78	Sequence 78, Appl
33	131.5	19.1	369	12	US-10-210-172-80	Sequence 80, Appl
34	131.5	19.1	382	12	US-10-315-664-93	Sequence 93, Appl
35	131.5	19.1	382	12	US-09-978-360A-425	Sequence 425, App
36	130	18.9	127	8	US-08-849-303-19	Sequence 19, Appl
37	130	18.9	127	16	US-10-655-136-19	Sequence 19, Appl
38	129.5	18.9	140	14	US-10-376-564-46	Sequence 46, Appl
39	129.5	18.9	140	14	US-10-376-564-48	Sequence 48, Appl
40	128	18.6	111	8	US-08-849-303-26	Sequence 26, Appl
41	128	18.6	111	16	US-10-655-136-26	Sequence 26, Appl
42	127.5	18.6	140	8	US-08-849-303-18	Sequence 18, Appl
43	127.5	18.6	140	16	US-10-655-136-18	Sequence 18, Appl
44	124.5	18.1	121	9	US-09-775-932-8	Sequence 8, Appli
45	124.5	18.1	141	8	US-08-849-303-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-10-162-335-82
; Sequence 82, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04

```
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 82
;    LENGTH: 390
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-162-335-82

  Query Match             90.0%;  Score 618;  DB 15;  Length 390;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-59;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 275

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 331


RESULT 2
US-10-162-335-70
; Sequence 70, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 70
;    LENGTH: 398
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-162-335-70

  Query Match             90.0%;  Score 618;  DB 15;  Length 398;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-59;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 339


RESULT 3
US-09-919-039-29
; Sequence 29, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
;  APPLICANT: Kaser, Matthew R.
;  TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
;  FILE REFERENCE: PA-0035 US
;  CURRENT APPLICATION NUMBER: US/09/919,039
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 60/222,113
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 401
;  SOFTWARE: PERL Program
; SEQ ID NO 29
;    LENGTH: 427
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20030108871A1 167507CD1
US-09-919-039-29

  Query Match             90.0%;  Score 618;  DB 10;  Length 427;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-59;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368
```

```
RESULT 4
US-10-162-335-72
; Sequence 72, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 72
;   LENGTH: 615
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-335-72

  Query Match             90.0%;  Score 618;  DB 15;  Length 615;
  Best Local Similarity   100.0%;  Pred. No. 1e-58;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 339


RESULT 5
US-10-162-335-74
; Sequence 74, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 74
;   LENGTH: 644
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-335-74

  Query Match             90.0%;  Score 618;  DB 15;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-58;
  Matches  116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312
```

```
Qy          63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368
```

RESULT 6
US-10-162-335-84
; Sequence 84, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 84
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-84

Query Match 90.0%; Score 618; DB 15; Length 644;
Best Local Similarity 100.0%; Pred. No. 1.1e-58;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy          63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 368
```

RESULT 7
US-10-316-253-217
; Sequence 217, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-217

Query Match 55.5%; Score 381; DB 14; Length 424;
Best Local Similarity 62.1%; Pred. No. 5.8e-33;
Matches 72; Conservative 14; Mismatches 30; Indels 0; Gaps 0;

```
Qy           3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
               | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||| ||||  |||
Db         246 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 305

Qy          63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
               ||  | |:|:|||| |||:|  |||  |||| ||||||||| ||: ||| |: |||
Db         306 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 361
```

RESULT 8
US-10-316-253-215
; Sequence 215, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 215
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-215

Query Match 55.3%; Score 380; DB 14; Length 430;
Best Local Similarity 62.1%; Pred. No. 7.7e-33;
Matches 72; Conservative 14; Mismatches 30; Indels 0; Gaps 0;

```
Qy          3 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 62
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||  ||||   |||
Db        252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         63 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTV 118
              ||   | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: |||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTV 367


RESULT 9
US-10-329-428-2
; Sequence 2, Application US/10329428
; Publication No. US20030114646A1
; GENERAL INFORMATION:
;  APPLICANT: Li, et al.
;  TITLE OF INVENTION: Human Cystatin F
;  FILE REFERENCE: PF265P1D2
;  CURRENT APPLICATION NUMBER: US/10/329,428
;  CURRENT FILING DATE:  2002-12-27
;  PRIOR APPLICATION NUMBER: 60/014,795
;  PRIOR FILING DATE: 1996-04-03
;  PRIOR APPLICATION NUMBER: 08/832,535
;  PRIOR FILING DATE: 1997-04-03
;  PRIOR APPLICATION NUMBER: 09/019,485
;  PRIOR FILING DATE: 1998-01-29
;  PRIOR APPLICATION NUMBER: 09/528,436
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-329-428-2

  Query Match             23.8%;  Score 163.5;  DB 14;  Length 145;
  Best Local Similarity   31.6%;  Pred. No. 9.7e-10;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |  :  | : |:      | |:| | :||:|||          : |:|
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 138


RESULT 10
US-09-746-783-197
; Sequence 197, Application US/09746783
; Publication No. US20030044935A1
;  GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John M.
;                  LaVallie, Edward R.
;                  Racie, Lisa A.
;                  Treacy, Maurice
;                  Spaulding, Vikki
;                  Agostino, Michael J.
;                  Howes, Steven H.
;                  Fechtel, Kim
;       TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
;                           ENCODING THEM
;       NUMBER OF SEQUENCES: 231
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: MA
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/746,783
;            FILING DATE: 21-Dec-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Milasincic, Debra J.
;            REGISTRATION NUMBER: 46,931
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 227-7400
;            TELEFAX: (617) 742-4214
;  INFORMATION FOR SEQ ID NO: 197:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 167 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 197:
US-09-746-783-197

  Query Match             23.8%;  Score 163.5;  DB 10;  Length 167;
  Best Local Similarity   31.6%;  Pred. No. 1.2e-09;
  Matches   37;  Conservative   22;  Mismatches   45;  Indels   13;  Gaps    3;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWE 124
                 ||| |  :  | : |:      | |:| | :||:|||          : |:|
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW---------LQHFE 160


RESULT 11
US-09-969-834-1
; Sequence 1, Application US/09969834
; Patent No. US20020102711A1
;  GENERAL INFORMATION:
;       APPLICANT: Bandman, Olga
;                  Goli, Surya K.
;       TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
;                           PROTEIN
;       NUMBER OF SEQUENCES: 4
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Drive
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/969,834
;            FILING DATE: 01-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/471,765
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US/08/791,522
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 09/471,765
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
```

```
;            REFERENCE/DOCKET NUMBER: PF-0193 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-855-0555
;            TELEFAX: 415-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 178 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            CLONE: 30443
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-969-834-1

  Query Match            23.7%;  Score 163;  DB 9;  Length 178;
  Best Local Similarity  34.0%;  Pred. No. 1.4e-09;
  Matches   35;  Conservative   20;  Mismatches   44;  Indels    4;  Gaps    2;

Qy         11 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 70
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         71 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPW 110
                 ||| |  :  | : |:       | |:| | :||:|||
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW 155


RESULT 12
US-09-775-932-14
; Sequence 14, Application US/09775932
; Patent No. US20020137671A1
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 121
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-14

  Query Match            20.2%;  Score 138.5;  DB 9;  Length 121;
  Best Local Similarity  31.5%;  Pred. No. 4.2e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          2 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 61

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         62 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 108


RESULT 13
US-09-775-932-12
; Sequence 12, Application US/09775932
; Patent No. US20020137671A1
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-12

  Query Match            20.2%;  Score 138.5;  DB 9;  Length 128;
  Best Local Similarity  31.5%;  Pred. No. 4.5e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          9 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 68

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         69 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 115


RESULT 14
US-09-940-497-2
; Sequence 2, Application US/09940497
; Patent No. US20020052476A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: Human Cystatin E
;  FILE REFERENCE: PF202P1D2
;  CURRENT APPLICATION NUMBER: US/09/940,497
;  CURRENT FILING DATE:  2001-08-29
;  PRIOR APPLICATION NUMBER: US 09/241,376
;  PRIOR FILING DATE: 1999-02-02
;  PRIOR APPLICATION NUMBER: US 08/744,138
;  PRIOR FILING DATE: 1996-11-05
;  PRIOR APPLICATION NUMBER: US 08/461,030
;  PRIOR FILING DATE: 1995-06-05
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-940-497-2

  Query Match            20.2%;  Score 138.5;  DB 9;  Length 149;
  Best Local Similarity  31.5%;  Pred. No. 5.5e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy         10 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 69
              | :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db         30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy         70 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 111
                 | | |       : :|| :|       |   | |: || ||||:
Db         90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 15
US-08-849-303-16
; Sequence 16, Application US/08849303
; Publication No. US20030221209A1
;  GENERAL INFORMATION:
;    APPLICANT:  Atkinson, Howard J.
;    APPLICANT:  McPherson, Michael J.
;    APPLICANT:  Urwin, Peter E.
;    TITLE OF INVENTION:  MODIFIED PROTEINASE INHIBITORS
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Klauber & Jackson
;       STREET:  411 Hackensack Avenue, 4th Floor
;       CITY:  Hackensack
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/849,303
;       FILING DATE:  21-MAY-1997
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  1321-1-003
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201-343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  112 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
US-08-849-303-16

  Query Match             20.0%;  Score 137.5;  DB 8;  Length 112;
  Best Local Similarity   28.8%;  Pred. No. 4.9e-07;
  Matches   32;  Conservative   25;  Mismatches   35;  Indels   19;  Gaps    4;

Qy          24 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 83
               |   ::| |:  ::: |  :|  :  ::  | :|| |||:|  ||:|     |||:|  :
Db          12 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 69

Qy          84 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVTVNHWECE 126
               :   :||        : :||      |: :||||||      |: :  : |:
Db          70 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 111


Search completed: September 24, 2004, 14:13:04
Job time : 45.704 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:07:01 ; Search time 14.268 Seconds
(without alignments)
445.051 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	600	90.9	117	1	US-08-193-114B-1	Sequence 1, Appli
2	594	90.0	117	5	PCT-US92-06809-1	Sequence 1, Appli
3	169	25.6	178	2	US-08-791-522-1	Sequence 1, Appli
4	169	25.6	178	3	US-09-314-777-1	Sequence 1, Appli
5	165.5	25.1	145	2	US-08-832-535-2	Sequence 2, Appli
6	165.5	25.1	145	3	US-09-019-485-2	Sequence 2, Appli
7	165.5	25.1	145	3	US-09-019-485-3	Sequence 3, Appli
8	165.5	25.1	145	3	US-09-431-480-9	Sequence 9, Appli
9	165.5	25.1	145	3	US-09-617-302-9	Sequence 9, Appli
10	165.5	25.1	145	4	US-09-528-436B-2	Sequence 2, Appli
11	155	23.5	27	3	US-08-676-242-11	Sequence 11, Appl
12	145	22.0	64	3	US-08-676-242-22	Sequence 22, Appl
13	145	22.0	64	3	US-09-402-732-1	Sequence 1, Appli
14	138.5	21.0	121	4	US-09-775-932-14	Sequence 14, Appl
15	138.5	21.0	128	4	US-09-775-932-12	Sequence 12, Appl
16	138.5	21.0	149	2	US-08-461-030C-2	Sequence 2, Appli
17	138.5	21.0	149	3	US-08-744-138-2	Sequence 2, Appli
18	138.5	21.0	149	3	US-09-431-480-8	Sequence 8, Appli
19	138.5	21.0	149	3	US-09-431-480-10	Sequence 10, Appl
20	138.5	21.0	149	3	US-09-617-302-8	Sequence 8, Appli
21	138.5	21.0	149	3	US-09-617-302-10	Sequence 10, Appl
22	138.5	21.0	149	4	US-09-241-376-2	Sequence 2, Appli
23	138.5	21.0	149	4	US-09-940-497-2	Sequence 2, Appli
24	136	20.6	112	4	US-08-849-303-16	Sequence 16, Appl
25	135	20.5	118	4	US-09-775-932-24	Sequence 24, Appl
26	134	20.3	148	5	PCT-US95-07135-2	Sequence 2, Appli
27	132	20.0	26	3	US-08-676-242-15	Sequence 15, Appl
28	130	19.7	127	4	US-08-849-303-19	Sequence 19, Appl
29	129.5	19.6	140	4	US-09-886-319A-46	Sequence 46, Appl
30	129.5	19.6	140	4	US-09-886-319A-48	Sequence 48, Appl
31	129.5	19.6	382	4	US-09-599-360B-93	Sequence 93, Appl
32	128.5	19.5	146	6	5432264-6	Patent No. 5432264
33	128	19.4	111	4	US-08-849-303-26	Sequence 26, Appl
34	127.5	19.3	120	4	US-09-775-932-2	Sequence 2, Appli
35	127.5	19.3	140	3	US-09-431-480-5	Sequence 5, Appli
36	127.5	19.3	140	3	US-09-617-302-5	Sequence 5, Appli
37	127.5	19.3	140	4	US-08-849-303-18	Sequence 18, Appl
38	127.5	19.3	145	2	US-08-832-535-11	Sequence 11, Appl
39	127.5	19.3	146	2	US-08-791-522-3	Sequence 3, Appli
40	127.5	19.3	146	3	US-08-744-138-3	Sequence 3, Appli
41	127.5	19.3	146	3	US-09-019-485-4	Sequence 4, Appli
42	127.5	19.3	146	3	US-09-314-777-3	Sequence 3, Appli
43	127.5	19.3	146	3	US-09-431-480-6	Sequence 6, Appli
44	127.5	19.3	146	3	US-09-617-302-6	Sequence 6, Appli
45	127.5	19.3	146	4	US-09-241-376-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-193-114B-1
; Sequence 1, Application US/08193114B
; Patent No. 5472945
; GENERAL INFORMATION:
; APPLICANT: Schmaier, Alvin H.
; APPLICANT: Jiang, Yongping
; TITLE OF INVENTION: Modulation of Blood
; TITLE OF INVENTION: Pressure and Inhibition of Platelet Activation
; TITLE OF INVENTION: with Kininogen Fragment
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna &
; ADDRESSEE: Monaco, P.C.
; STREET: 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,114B
; FILING DATE: 9 February 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Application
; APPLICATION NUMBER: Serial No. 5472945 07/744,545
; FILING DATE: 13 August 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-137 CI1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5472945e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: peptide
; TOPOLOGY: linear
US-08-193-114B-1

Query Match 90.9%; Score 600; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.8e-59;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          12 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 71
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 60

Qy          72 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 123
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 112


RESULT 2
PCT-US92-06809-1
; Sequence 1, Application PC/TUS9206809
;  GENERAL INFORMATION:
;    APPLICANT:  Schmaier, Alvin H.
;    APPLICANT:  Jiang, Yongping
;    TITLE OF INVENTION:  Modulation of Blood
;    TITLE OF INVENTION:  Pressure by Altering Bradykinin Levels
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of the
;      ADDRESSEE:  Commonwealth System of Higher Education
;      STREET:  406 University Services
;      STREET:  Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06809
;      FILING DATE:  19910813
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. Application
;      APPLICATION NUMBER:  Serial No.  744,545
;      FILING DATE:  13 August 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-137
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-06809-1

  Query Match             90.0%;  Score 594;  DB 5;  Length 117;
  Best Local Similarity   98.2%;  Pred. No. 2.7e-58;
  Matches  110;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          12 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 71
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 60

Qy          72 RETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 123
               |||||||||||||||||||||:|||||||||||||||||||||:||||
Db          61 RETTCSKESNEELTESCETKKLGESLDCNAEVYVVPWEKKIYPTVNCEPLGM 112


RESULT 3
US-08-791-522-1
; Sequence 1, Application US/08791522
; Patent No. 5935817
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/791,522
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-08-791-522-1

  Query Match             25.6%;  Score 169;  DB 2;  Length 178;
  Best Local Similarity   32.5%;  Pred. No. 5.4e-11;
  Matches   39;  Conservative   22;  Mismatches   49;  Indels   10;  Gaps    4;

Qy           9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
               :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db          54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVN--CQPLGM 123
                  ||| |  :  | : |:       | |:| | :||:||||     | :   | || :
Db         114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW----VPALRGACSPLSL 168


RESULT 4
US-09-314-777-1
; Sequence 1, Application US/09314777
; Patent No. 6110686
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN CYSTATIN-LIKE
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/314,777
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/791,522
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0193 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  178 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  30443
US-09-314-777-1

  Query Match             25.6%;  Score 169;  DB 3;  Length 178;
  Best Local Similarity   32.5%;  Pred. No. 5.4e-11;
  Matches   39;  Conservative   22;  Mismatches   49;  Indels   10;  Gaps    4;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVN--CQPLGM 123
                ||| |  :  | : |:      | |:| | :||:||||     | :   | || :
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW----VPALRGACSPLSL 168


RESULT 5
US-08-832-535-2
; Sequence 2, Application US/08832535
; Patent No. 5919658
;  GENERAL INFORMATION:
;    APPLICANT:  NI, JIAN
;    APPLICANT:  LI, HAODONG
;    APPLICANT:  YU, GUO-LIANG
;    APPLICANT:  GENTZ, REINER L
;    TITLE OF INVENTION:  HUMAN CYSTATIN F
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/832,535
;      FILING DATE:  03-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KIMBALL, PAUL C.
;      REGISTRATION NUMBER:  34,610
;      REFERENCE/DOCKET NUMBER:  PF265
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-832-535-2

  Query Match             25.1%;  Score 165.5;  DB 2;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                ||| |  :  | : |:      | |:| | :||:|||| :    | : |
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 6
US-09-019-485-2
; Sequence 2, Application US/09019485
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-2

  Query Match             25.1%;  Score 165.5;  DB 3;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91
```

```
Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
               |||  |   :  |  : |:       | |:| | :||:|||| :    |  : |
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 7
US-09-019-485-3
; Sequence 3, Application US/09019485
; Patent No. 6066617
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Haodong
;    APPLICANT:  Yu, Guo-Liang
;    APPLICANT:  Gentz, Reiner
;    APPLICANT:  Ni, Jian
;    TITLE OF INVENTION:  Cystatin F
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/019,485
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Robert H.
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PF265P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  3013098504
;      TELEFAX:  3013098439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  145 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-019-485-3

  Query Match             25.1%;  Score 165.5;  DB 3;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy           9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
               :::  | |: | || | : :   ::: | |    |  | ||    : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
               |||  |   :  |  : |:       | |:| | :||:|||| :    |  : |
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 8
US-09-431-480-9
; Sequence 9, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
;  APPLICANT: Holloway, James L.
;  APPLICANT: Feldhaus, Andrew
;  TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE: 98-72
;  CURRENT APPLICATION NUMBER: US/09/431,480
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: 60/109,217
;  EARLIER FILING DATE: 1998-11-20
;  EARLIER APPLICATION NUMBER: 60/156,382
;  EARLIER FILING DATE: 1999-09-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-431-480-9

  Query Match             25.1%;  Score 165.5;  DB 3;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy           9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
               :::  | |: | || | : :   ::: | |    |  | ||    : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
               |||  |   :  |  : |:       | |:| | :||:|||| :    |  : |
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 9
US-09-617-302-9
; Sequence 9, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
;  APPLICANT:  Holloway, James L.
;  APPLICANT:  Feldhaus, Andrew
;  TITLE OF INVENTION:  TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
;  FILE REFERENCE:  98-72 C1
;  CURRENT APPLICATION NUMBER: US/09/617,302
;  CURRENT FILING DATE:  2000-07-17
;  PRIOR APPLICATION NUMBER:  09/431,480
;  PRIOR FILING DATE:  1999-11-01
;  PRIOR APPLICATION NUMBER:  60/109,217
;  PRIOR FILING DATE:  1998-11-20
;  PRIOR APPLICATION NUMBER:  60/156,382
;  PRIOR FILING DATE:  1999-09-28
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-617-302-9

  Query Match             25.1%;  Score 165.5;  DB 3;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy           9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
               :::  | |: | || | : :   ::: | |    |  | ||    : :| ||:| | || ::
Db          32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
               |||  |   :  |  : |:       | |:| | :||:|||| :    |  : |
Db          92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 10
US-09-528-436B-2
; Sequence 2, Application US/09528436B
; Patent No. 6576745
; GENERAL INFORMATION:
;  APPLICANT: Li, et al.
;  TITLE OF INVENTION: Human Cystatin F
;  FILE REFERENCE: PF265P1D1
;  CURRENT APPLICATION NUMBER: US/09/528,436B
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER: 09/019,485
```

```
;  PRIOR FILING DATE: 1998-01-29
;  PRIOR APPLICATION NUMBER: 08/832,535
;  PRIOR FILING DATE: 1999-04-03
;  PRIOR APPLICATION NUMBER: 60/014,795
;  PRIOR FILING DATE: 1996-04-03
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-528-436B-2

  Query Match             25.1%;  Score 165.5;  DB 4;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 1e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                ||| |  :  | : |:      | |:| | :||:|||| :    | : |
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 11
US-08-676-242-11
; Sequence 11, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 27
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Bradykinin
;   OTHER INFORMATION: analog
US-08-676-242-11

  Query Match             23.5%;  Score 155;  DB 3;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-10;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         97 LDCNAEVYVVPWEKKIYPTVNCQPLGM 123
              |||||||||||||||||||||||||||
Db          1 LDCNAEVYVVPWEKKIYPTVNCQPLGM 27


RESULT 12
US-08-676-242-22
; Sequence 22, Application US/08676242C
; Patent No. 6143719
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of Michigan
;  APPLICANT: Schmaier, Alvin H.
;  APPLICANT: Hasan, Ahmed A.K.
;  TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors
;  FILE REFERENCE: 8820-2 US
;  CURRENT APPLICATION NUMBER: US/08/676,242C
;  CURRENT FILING DATE: 2000-07-16
;  EARLIER APPLICATION NUMBER: 60/000,096
;  EARLIER FILING DATE: 1995-06-09
;  EARLIER APPLICATION NUMBER: PCT/US96/09940
;  EARLIER FILING DATE: 1996-06-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 64
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: From human
;   OTHER INFORMATION: kininogen heavy chain
US-08-676-242-22

  Query Match             22.0%;  Score 145;  DB 3;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         99 CNAEVYVVPWEKKIYPTVNCQPLGM 123
              |||||||||||||||||||||||||
Db          1 CNAEVYVVPWEKKIYPTVNCQPLGM 25


RESULT 13
US-09-402-732-1
; Sequence 1, Application US/09402732
; Patent No. 6251855
; GENERAL INFORMATION:
;  APPLICANT:  Schmainer, Alvin H.
;  APPLICANT:  Hasan, A.K. Ahmed
;  TITLE OF INVENTION:  Bradykinin Analogs As Selective Inhibitors of Cell
;  TITLE OF INVENTION:  Activation
;  FILE REFERENCE:  8820-3
;  CURRENT APPLICATION NUMBER: US/09/402,732
;  CURRENT FILING DATE:  1999-12-01
;  PRIOR APPLICATION NUMBER:  60/046,085
;  PRIOR FILING DATE:  1997-04-23
;  PRIOR APPLICATION NUMBER:  PCT/US98/08015
;  PRIOR FILING DATE:  1998-04-21
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 64
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Segment of
;   OTHER INFORMATION: human kininogen (residues 333-396 threreof)
US-09-402-732-1

  Query Match             22.0%;  Score 145;  DB 3;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         99 CNAEVYVVPWEKKIYPTVNCQPLGM 123
              |||||||||||||||||||||||||
Db          1 CNAEVYVVPWEKKIYPTVNCQPLGM 25


RESULT 14
US-09-775-932-14
; Sequence 14, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
```

```
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 121
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-14

  Query Match            21.0%;  Score 138.5;  DB 4;  Length 121;
  Best Local Similarity  31.5%;  Pred. No. 7.7e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
              |  :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          2 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 61

Qy         68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                 | | |       : :|| :|       |   | |: || ||||:
Db         62 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 108


RESULT 15
US-09-775-932-12
; Sequence 12, Application US/09775932
; Patent No. 6534477
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-12

  Query Match            21.0%;  Score 138.5;  DB 4;  Length 128;
  Best Local Similarity  31.5%;  Pred. No. 8.3e-08;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
              |  :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db          9 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 68

Qy         68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                 | | |       : :|| :|       |   | |: || ||||:
Db         69 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 115


Search completed: September 24, 2004, 14:11:36
Job time : 27.268 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:02:27 ; Search time 49.692 Seconds
(without alignments)
699.375 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	123	3	AAY95426	Aay95426 Human hig
2	660	100.0	304	6	ABP70801	Abp70801 Human ext
3	660	100.0	322	6	ABP70799	Abp70799 Human ext
4	660	100.0	329	6	ABU92044	Abu92044 Human pro
5	660	100.0	358	6	ABP70800	Abp70800 Human ext
6	660	100.0	390	6	ABU99149	Abu99149 Novel hum
7	660	100.0	398	6	ABU99143	Abu99143 Novel hum
8	660	100.0	427	8	ADE76864	Ade76864 Human pro
9	660	100.0	615	6	ABU99144	Abu99144 Novel hum
10	660	100.0	626	5	ABB78707	Abb78707 Human hig
11	660	100.0	644	4	ABG21101	Abg21101 Novel hum
12	660	100.0	644	5	ABB78710	Abb78710 Human hig
13	660	100.0	644	6	ABU99150	Abu99150 Novel hum
14	660	100.0	644	6	ABU99145	Abu99145 Novel hum
15	628	95.2	122	3	AAB37447	Aab37447 Human kin
16	594	90.0	117	2	AAR33350	Aar33350 Domaine 3
17	589.5	89.3	435	4	ABG21105	Abg21105 Novel hum
18	477	72.3	436	1	AAP40257	Aap40257 Bradykini
19	453	68.6	357	6	ABR41202	Abr41202 Human DIT
20	450	68.2	434	1	AAP40633	Aap40633 Bradykini
21	409	62.0	235	5	ABG60077	Abg60077 Human DIT
22	320.5	48.6	248	4	ABG21102	Abg21102 Novel hum
23	316	47.9	369	4	ABG21099	Abg21099 Novel hum
24	190	28.8	305	4	ABG21100	Abg21100 Novel hum
25	175	26.5	167	2	AAW98907	Aaw98907 Mouse IMC
26	169	25.6	178	2	AAW69734	Aaw69734 Human cys
27	166	25.2	32	3	AAY95418	Aay95418 Anti-angi
28	165.5	25.1	126	3	AAB37445	Aab37445 Human cys
29	165.5	25.1	145	2	AAW32323	Aaw32323 Mature hu
30	165.5	25.1	145	2	AAW31902	Aaw31902 Human cys
31	165.5	25.1	145	2	AAY25708	Aay25708 Human cys
32	165.5	25.1	145	4	AAE02410	Aae02410 Human cys
33	165.5	25.1	145	4	AAE04439	Aae04439 Human cys
34	165.5	25.1	145	7	ADD14098	Add14098 Human src
35	165.5	25.1	167	2	AAY02287	Aay02287 Secreted
36	165.5	25.1	167	7	ADA45154	Ada45154 Human pol
37	161	24.4	32	3	AAY95408	Aay95408 Anti-angi
38	157	23.8	122	3	AAB37446	Aab37446 Human kin
39	156.5	23.7	167	2	AAW98910	Aaw98910 Mouse IMC
40	155	23.5	27	2	AAW54335	Aaw54335 Bradykini
41	153	23.2	27	3	AAY95425	Aay95425 Anti-angi
42	145	22.0	64	2	AAW54341	Aaw54341 Bradykini
43	145	22.0	64	2	AAW77418	Aaw77418 Kininogen
44	144.5	21.9	121	3	AAY81200	Aay81200 Human mut
45	144.5	21.9	128	3	AAY81189	Aay81189 Human mut

ALIGNMENTS

RESULT 1
AAY95426
ID AAY95426 standard; peptide; 123 AA.
XX
AC AAY95426;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human high mol.wt. kininogen domain 3.
XX
KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic; therapy;
KW human; D3 peptide.
XX
OS Homo sapiens.
XX
PN WO200035407-A2.
XX
PD 22-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US028465.
XX
PR 16-DEC-1998; 98US-0112427P.
XX
PA (UTEM) UNIV TEMPLE.
PA (MCCR/) MCCRAE R K.
XX
PI Mccrae RK;
XX
DR WPI; 2000-442247/38.
XX
PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain 3
PT analog.
XX
PS Disclosure; Page 4; 44pp; English.
XX
CC The present sequence is that of domain 3 of human high mol.wt. kininogen
CC (HK). The invention provides peptides (see AAY95405-24) that are
CC analogues of certain sites in the HK domain 3, specifically Asn275-
CC Lys282, Cys246-Cys249, Leu331-Tyr338 and Tyr299-Ser314. The peptides, in
CC which native Cys residues may be replaced by Ala residues, inhibit
CC endothelial cell proliferation and may also induce endothelial cell
CC apoptosis. Compositions including the peptides are used in claimed
CC methods for inhibiting angiogenesis, inhibiting endothelial cell
CC proliferation, and inducing endothelial cell apoptosis. Cancer,

CC rheumatoid arthritis, and ocular disorders characterized by undesired
CC vascularization of the retina are treated
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 660; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.3e-66;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120

Qy        121 LGM 123
              |||
Db        121 LGM 123
```

RESULT 2
ABP70801
ID ABP70801 standard; protein; 304 AA.
XX
AC ABP70801;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extracellular messenger, EXMES-28.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42388.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX
PS Claim 1; Page 207; 224pp; English.
XX
CC The present invention relates to novel human extracellular messenger
CC proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC for diagnosing or treating a disease or condition associated with
CC decreased expression or overexpression of functional EXMES e.g.
CC autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC cancer
XX
SQ Sequence 304 AA;

Query Match 100.0%; Score 660; DB 6; Length 304;
Best Local Similarity 100.0%; Pred. No. 8e-66;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 189

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 249

Qy        121 LGM 123
              |||
Db        250 LGM 252
```

RESULT 3
ABP70799
ID ABP70799 standard; protein; 322 AA.
XX
AC ABP70799;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extracellular messenger, EXMES-26.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42386.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX
PS Claim 1; Page 205-206; 224pp; English.
XX
CC The present invention relates to novel human extracellular messenger

CC proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC for diagnosing or treating a disease or condition associated with
CC decreased expression or overexpression of functional EXMES e.g.
CC autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC cancer
XX
SQ Sequence 322 AA;

Query Match 100.0%; Score 660; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.7e-66;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 207

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 267

Qy        121 LGM 123
              |||
Db        268 LGM 270
```

RESULT 4
ABU92044
ID ABU92044 standard; protein; 329 AA.
XX
AC ABU92044;
XX
DT 15-JUL-2003 (first entry)
XX
DE Human protein modification and maintenance molecule-24 (PMMM-24).
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR N-PSDB; ACA92439.
XX
PT New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 1; Page 264-265; 311pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention
XX
SQ Sequence 329 AA;

Query Match 100.0%; Score 660; DB 6; Length 329;
Best Local Similarity 100.0%; Pred. No. 9e-66;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 214

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 274

Qy        121 LGM 123
              |||
Db        275 LGM 277
```

RESULT 5
ABP70800
ID ABP70800 standard; protein; 358 AA.
XX
AC ABP70800;
XX
DT 26-AUG-2003 (first entry)
XX
DE Human extracellular messenger, EXMES-27.
XX
KW Human; extracellular messenger; EXMES; cytostatic; antidiabetic;
KW immunosuppressive; autoimmune disorder; inflammatory disorder; diabetes;
KW endocrine disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO2003018612-A2.
XX
PD 06-MAR-2003.
XX
PF 22-AUG-2002; 2002WO-US027213.
XX
PR 24-AUG-2001; 2001US-0314811P.
PR 14-DEC-2001; 2001US-0340584P.
PR 18-JAN-2002; 2002US-0350595P.
PR 11-MAR-2002; 2002US-0363432P.
PR 15-MAR-2002; 2002US-0364607P.
PR 05-APR-2002; 2002US-0370761P.
PR 24-JUN-2002; 2002US-0391378P.
XX

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Duggan BM, Lee S, Baughn MR, Hafalia AJA, Walia NK, Elliott VS;
PI Lee SY, Khare R, Wilson AD, Jin P, Hawkins PR, Tang YT, Burford N;
PI Ding L, Yao MG, Becha SD, Tran UK, Chien D, Zebarjadian Y;
PI Richardson TW, Kable AE, Chang H, Swarnakar A, Lehr-Mason PM;
PI Ramkumar J;
XX
DR WPI; 2003-278643/27.
DR N-PSDB; ACC42387.
XX
PT New human extracellular messenger (EXMES) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional EXMES e.g. autoimmune
PT disorders or cancer.
XX
PS Claim 1; Page 206; 224pp; English.
XX
CC The present invention relates to novel human extracellular messenger
CC proteins (EXMES-1 to-28; ABP70774-ABP70801) and their coding sequences
CC (ACC42361-ACC42388). The proteins are useful for preparing a composition
CC for diagnosing or treating a disease or condition associated with
CC decreased expression or overexpression of functional EXMES e.g.
CC autoimmune/inflammatory disorders, diabetes, endocrine disorders or
CC cancer
XX
SQ Sequence 358 AA;

Query Match 100.0%; Score 660; DB 6; Length 358;
Best Local Similarity 100.0%; Pred. No. 1e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
Db 184 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 243

Qy 61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
Db 244 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 303

Qy 121 LGM 123
Db 304 LGM 306

RESULT 6
ABU99149
ID ABU99149 standard; protein; 390 AA.
XX
AC ABU99149;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human GPCR related protein NOV12g.
XX
KW Human; G-protein coupled receptor related protein; GPCR related protein;
KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03653.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 147; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 390 AA;

Query Match 100.0%; Score 660; DB 6; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.1e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
Db 216 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 275

```
Qy          61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         276 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 335

Qy         121 LGM 123
               |||
Db         336 LGM 338
```

RESULT 7
ABU99143
ID ABU99143 standard; protein; 398 AA.
XX
AC ABU99143;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human GPCR related protein NOV12a.
XX
KW Human; G-protein coupled receptor related protein; GPCR related protein;
KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03647.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 143; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 398 AA;

Query Match 100.0%; Score 660; DB 6; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.2e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy          61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 343

Qy         121 LGM 123
               |||
Db         344 LGM 346
```

RESULT 8
ADE76864
ID ADE76864 standard; protein; 427 AA.
XX
AC ADE76864;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein expressed in a liver disorder #9.
XX
KW human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
KW high-throughput screening; differential gene expression; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003108871-A1.
XX
PD 12-JUN-2003.
XX
PF 30-JUL-2001; 2001US-00919039.
XX
PR 28-JUL-2000; 2000US-0222113P.
XX
PA (KASE/) KASER M R.
XX
PI Kaser MR;

XX
DR WPI; 2004-031227/03.
DR N-PSDB; ADE76863.
XX
PT Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX
PS Claim 1; SEQ ID NO 29; 41pp; English.
XX
CC The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX
SQ Sequence 427 AA;

Query Match 100.0%; Score 660; DB 8; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.3e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy      61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy     121 LGM 123
          |||
Db     373 LGM 375
```

RESULT 9
ABU99144
ID ABU99144 standard; protein; 615 AA.
XX
AC ABU99144;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human GPCR related protein NOV12b.
XX
KW Human; G-protein coupled receptor related protein; GPCR related protein;
KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03648.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 144; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 615 AA;

Query Match 100.0%; Score 660; DB 6; Length 615;
Best Local Similarity 100.0%; Pred. No. 2.1e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 343

Qy        121 LGM 123
              |||
Db        344 LGM 346
```

```
RESULT 10
ABB78707
ID   ABB78707 standard; protein; 626 AA.
XX
AC   ABB78707;
XX
DT   18-JUL-2002  (first entry)
XX
DE   Human high molecular weight kininogen (HK) mature protein SEQ ID NO:1.
XX
KW   Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW   D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumour;
KW   antiatherosclerotic; vasotropic; vulnerary; tranquilliser; thrombolytic;
KW   ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW   antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          384. .508
FT                   /label= D5_domain
XX
PN   WO200214369-A2.
XX
PD   21-FEB-2002.
XX
PF   24-JUL-2001; 2001WO-US023185.
XX
PR   24-JUL-2000; 2000US-0220194P.
XX
PA   (ATTE-) ATTENUON LLC.
XX
PI   Mazar AP,  Juarez JC;
XX
DR   WPI; 2002-393611/42.
XX
PT   Novel human kininogen D5 domain polypeptides useful for treating
PT   conditions associated with endothelial cell migration, proliferation,
PT   invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
PT   hyperplasia.
XX
PS   Disclosure; Page 13; 84pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) that
CC   corresponds to the D5 domain of human kininogen, or biologically active
CC   peptide fragment, homologue or functional derivative, and which: (a)
CC   inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
CC   cells (EC); (c) activates signalling pathways leading to the introduction
CC   of apoptosis in EC; and/or (d) inhibits the signalling pathway required
CC   for maintenance of EC viability. (I) has cytostatic, antitumour,
CC   antiatherosclerotic, vasotropic, vulnerary, tranquilliser, thrombolytic,
CC   ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
CC   antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
CC   specific for an epitope of (I) is useful for inhibiting tumour growth or
CC   angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
CC   or trimeric fusion polypeptide (III) can be used for inhibiting EC
CC   migration, proliferation, invasion, or angiogenesis, or for inducing EC
CC   apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
CC   comprising (I), (II), or (III), can be used for treating a subject having
CC   a disease or condition associated with undesired EC migration,
CC   proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
CC   for isolating a D5 domain binding molecule from a complex mixture and for
CC   isolating or enriching cells expressing D5 domain binding sites from a
CC   cell mixture. The present sequence represents the mature human high
CC   molecular weight kininogen (HK) protein, which is given in the
CC   exemplification of the present invention
XX
SQ   Sequence 626 AA;

  Query Match             100.0%;  Score 660;  DB 5;  Length 626;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-65;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 294

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 354

Qy        121 LGM 123
              |||
Db        355 LGM 357
```

```
RESULT 11
ABG21101
ID   ABG21101 standard; protein; 644 AA.
XX
AC   ABG21101;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #21092.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS85288.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 51460; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
```

```
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 644 AA;

  Query Match             100.0%;  Score 660;  DB 4;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-65;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 12
ABB78710
ID   ABB78710 standard; protein; 644 AA.
XX
AC   ABB78710;
XX
DT   18-JUL-2002  (first entry)
XX
DE   Human high molecular weight kininogen (HK) protein.
XX
KW   Human; kininogen; high molecular weight kininogen; HK; D5 domain;
KW   D5 receptor; angiogenesis; endothelial cell; cytostatic; antitumour;
KW   antiatherosclerotic; vasotropic; vulnerary; tranquilliser; thrombolytic;
KW   ophthalmological; gynaecological; antiulcer; antidiabetic; antiarthritic;
KW   antiangiogenic; antiapoptotic; endocrine; apoptosis; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= signal
FT   Protein         19. .644
FT                   /label= mature_human_high_molecular_weight_kininogen
FT   Disulfide-bond  28. .614
FT   Disulfide-bond  83. .94
FT   Disulfide-bond  107. .126
FT   Disulfide-bond  142. .145
FT   Disulfide-bond  206. .218
FT   Disulfide-bond  229. .248
FT   Disulfide-bond  264. .267
FT   Disulfide-bond  328. .340
FT   Disulfide-bond  351. .370
FT   Domain          402. .526
FT                   /label= D5_domain
XX
PN   WO200214369-A2.
XX
PD   21-FEB-2002.
XX
PF   24-JUL-2001; 2001WO-US023185.
XX
PR   24-JUL-2000; 2000US-0220194P.
XX
PA   (ATTE-) ATTENUON LLC.
XX
PI   Mazar AP,  Juarez JC;
XX
DR   WPI; 2002-393611/42.
XX
PT   Novel human kininogen D5 domain polypeptides useful for treating
PT   conditions associated with endothelial cell migration, proliferation,
PT   invasion or angiogenesis, e.g. arthritis, macular degeneration, benign
PT   hyperplasia.
XX
PS   Disclosure; Fig 1B-E; 84pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) that
CC   corresponds to the D5 domain of human kininogen, or biologically active
CC   peptide fragment, homologue or functional derivative, and which: (a)
CC   inhibits angiogenesis; (b) binds to the D5 binding site on endothelial
CC   cells (EC); (c) activates signalling pathways leading to the introduction
CC   of apoptosis in EC; and/or (d) inhibits the signalling pathway required
CC   for maintenance of EC viability. (I) has cytostatic, antitumour,
CC   antiatherosclerotic, vasotropic, vulnerary, tranquilliser, thrombolytic,
CC   ophthalmological, gynaecological, antiulcer, antidiabetic, antiarthritic,
CC   antiangiogenic, antiapoptotic and endocrine activities. An antibody (IX)
CC   specific for an epitope of (I) is useful for inhibiting tumour growth or
CC   angiogenesis in a subject. (I), a D5 fusion polypeptide (II) or a dimeric
CC   or trimeric fusion polypeptide (III) can be used for inhibiting EC
CC   migration, proliferation, invasion, or angiogenesis, or for inducing EC
CC   apoptosis. An angiogenic EC-targeting pharmaceutical composition (X)
CC   comprising (I), (II), or (III), can be used for treating a subject having
CC   a disease or condition associated with undesired EC migration,
CC   proliferation, invasion or angiogenesis. (I), (II), or (III) can be used
CC   for isolating a D5 domain binding molecule from a complex mixture and for
CC   isolating or enriching cells expressing D5 domain binding sites from a
CC   cell mixture. The present sequence represents the human high molecular
CC   weight kininogen (HK) protein, which is given in the exemplification of
CC   the present invention
XX
SQ   Sequence 644 AA;

  Query Match             100.0%;  Score 660;  DB 5;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-65;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 13
ABU99150
ID   ABU99150 standard; protein; 644 AA.
XX
AC   ABU99150;
XX
DT   01-AUG-2003  (first entry)
XX
DE   Novel human GPCR related protein NOV12h.
XX
KW   Human; G-protein coupled receptor related protein; GPCR related protein;
KW   NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW   immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW   antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW   NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
```

KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03654.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 148; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 644 AA;

Query Match 100.0%; Score 660; DB 6; Length 644;
Best Local Similarity 100.0%; Pred. No. 2.3e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375
```

RESULT 14
ABU99145
ID ABU99145 standard; protein; 644 AA.
XX
AC ABU99145;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human GPCR related protein NOV12c.
XX
KW Human; G-protein coupled receptor related protein; GPCR related protein;
KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
XX
OS Homo sapiens.
XX
PN WO200299116-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US017428.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298556P.
PR 21-JUN-2001; 2001US-0299949P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301550P.
PR 13-AUG-2001; 2001US-0311972P.
PR 27-AUG-2001; 2001US-0315071P.
PR 29-AUG-2001; 2001US-0315660P.
PR 14-SEP-2001; 2001US-0322293P.
PR 17-SEP-2001; 2001US-0322706P.
PR 14-DEC-2001; 2001US-0341186P.
PR 28-FEB-2002; 2002US-0361189P.
PR 12-MAR-2002; 2002US-0363673P.
PR 12-MAR-2002; 2002US-0363676P.
PR 03-JUN-2002; 2002US-00363676.
XX
PA (CURA-) CURAGEN CORP.

XX
PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
PI Pena CEA, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;
PI Voss EZ, Zerhusen BD;
XX
DR WPI; 2003-140627/13.
DR N-PSDB; ACD03649.
XX
PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 144-145; 332pp; English.
XX
CC The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV
XX
SQ Sequence 644 AA;

Query Match 100.0%; Score 660; DB 6; Length 644;
Best Local Similarity 100.0%; Pred. No. 2.3e-65;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375
```

RESULT 15
AAB37447
ID AAB37447 standard; protein; 122 AA.
XX
AC AAB37447;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human kininogen D3.
XX
KW Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
XX
OS Homo sapiens.
XX
PN WO200064945-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB001571.
XX
PR 22-APR-1999; 99GB-00009133.
XX
PA (BABR-) BABRAHAM INST.
XX
PI Abrahamson M, Barrett AJ;
XX
DR WPI; 2000-687316/67.
XX
PT Inhibition of mammalian legumain or legumain-related endopeptidase by
PT cystatin involves interaction with second papain-non-reactive site of
PT cystatin.
XX
PS Disclosure; Fig 4; 45pp; English.
XX
CC The present invention relates to inhibition of the enzymatic activity of
CC legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC involves an interaction between legumain and a papain-non-reactive site
CC of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC performs a protein-processing function. The present sequence is human
CC kininogen D3, which was used in the present invention. Kininogen is a
CC type 3 cystatin
XX
SQ Sequence 122 AA;

Query Match 95.2%; Score 628; DB 3; Length 122;
Best Local Similarity 100.0%; Pred. No. 9.5e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          7 PPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYF 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYF 60

Qy         67 IDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 123
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 117
```

Search completed: September 24, 2004, 14:08:37
Job time : 52.692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:04:32 ; Search time 8.364 Seconds
(without alignments)
765.738 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	644	1	KNG_HUMAN	P01042 homo sapien
2	477	72.3	436	1	KNL1_BOVIN	P01046 bos taurus
3	477	72.3	621	1	KNH1_BOVIN	P01044 bos taurus
4	450	68.2	434	1	KNL2_BOVIN	P01047 bos taurus
5	450	68.2	619	1	KNH2_BOVIN	P01045 bos taurus
6	430	65.2	661	1	KNG_MOUSE	O08677 mus musculu
7	426	64.5	639	1	KNG_RAT	P08934 rattus norv
8	409	62.0	430	1	KNT2_RAT	P08932 rattus norv
9	401	60.8	430	1	KNT1_RAT	P01048 rattus norv
10	175	26.5	144	1	CYTF_MOUSE	O89098 mus musculu
11	165.5	25.1	145	1	CYTF_HUMAN	O76096 homo sapien
12	138.5	21.0	149	1	CYTM_HUMAN	Q15828 homo sapien
13	136	20.6	148	1	CYTC_BOVIN	P01035 bos taurus
14	133.5	20.2	146	1	CYTC_MACMU	O19092 macaca mula
15	132.5	20.1	146	1	CYTC_SAISC	O19093 saimiri sci
16	130	19.7	127	1	CYTC_RAT	P14841 rattus norv
17	129.5	19.6	140	1	CYTC_MOUSE	P21460 mus musculu
18	129.5	19.6	378	1	FETB_RAT	Q9qx79 rattus norv
19	129.5	19.6	382	1	FETB_HUMAN	Q9ugm5 homo sapien
20	128	19.4	111	1	CYT_BITAR	P08935 bitis ariet
21	127.5	19.3	146	1	CYTC_HUMAN	P01034 homo sapien
22	125	18.9	141	1	CYTT_HUMAN	P09228 homo sapien
23	124.5	18.9	148	1	CYTC_RABIT	O97862 oryctolagus
24	122.5	18.6	116	1	CYT_COTJA	P81061 coturnix co
25	118.5	18.0	139	1	CYT_CHICK	P01038 gallus gall
26	115.5	17.5	388	1	FETB_MOUSE	Q9qxc1 mus musculu
27	113.5	17.2	122	1	CYTA_SARPE	P31727 sarcophaga
28	113	17.1	141	1	CYTS_RAT	P19313 rattus norv
29	109.5	16.6	141	1	CYTN_HUMAN	P01037 homo sapien
30	109	16.5	141	1	CYTS_HUMAN	P01036 homo sapien
31	108.5	16.4	130	1	CYT_ONCKE	Q98967 oncorhynchu
32	108	16.4	129	1	CYT_CYPCA	P35481 cyprinus ca
33	106.5	16.1	130	1	CYT_ONCMY	Q91195 oncorhynchu
34	105.5	16.0	162	1	CYTX_ONCVO	P22085 onchocerca
35	101	15.3	139	1	CS11_MOUSE	Q9d269 mus musculu
36	95	14.4	165	1	CSTL_HUMAN	Q9h114 homo sapien
37	94.5	14.3	142	1	CST8_MOUSE	P32766 mus musculu
38	93.5	14.2	137	1	CS11_HUMAN	Q9h112 homo sapien
39	93.5	14.2	345	1	A2HS_MOUSE	P29699 mus musculu
40	93	14.1	99	1	CYT_NAJAT	P81714 naja atra (
41	91.5	13.9	135	1	CYTI_MAIZE	P31726 zea mays (m
42	91	13.8	142	1	CYTD_HUMAN	P28325 homo sapien
43	90.5	13.7	352	1	A2HS_RAT	P24090 rattus norv
44	89.5	13.6	142	1	CST8_HUMAN	O60676 homo sapien
45	86.5	13.1	142	1	CST8_RAT	O88969 rattus norv

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042; P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens.";
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution.";
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RT its identity with low molecular weight kininogen.";
RL Biochemistry 23:5691-5697(1984).
RN [4]
RP SEQUENCE OF 379-644.
RX MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschen A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen.";
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens.";
RL Fed. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

```
RL   Seikagaku 56:808-808(1984).
RN   [7]
RP   CARBOHYDRATE-LINKAGE SITE ASN-294.
RX   MEDLINE=22660472; PubMed=12754519;
RA   Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT   "Identification and quantification of N-linked glycoproteins using
RT   hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL   Nat. Biotechnol. 21:660-666(2003).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
CC       clotting.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=P01042-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=P01042-2; Sequence=VSP_001261, VSP_001262;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02566; AAA35497.1; -.
DR   EMBL; M11437; AAB59550.1; -.
DR   EMBL; M11438; AAB59550.1; JOINED.
DR   EMBL; M11521; AAB59550.1; JOINED.
DR   EMBL; M11522; AAB59550.1; JOINED.
DR   EMBL; M11523; AAB59550.1; JOINED.
DR   EMBL; M11524; AAB59550.1; JOINED.
DR   EMBL; M11525; AAB59550.1; JOINED.
DR   EMBL; M11526; AAB59550.1; JOINED.
DR   EMBL; M11527; AAB59550.1; JOINED.
DR   EMBL; M11528; AAB59550.1; JOINED.
DR   EMBL; M11437; AAB59551.1; -.
DR   EMBL; M11438; AAB59551.1; JOINED.
DR   EMBL; M11521; AAB59551.1; JOINED.
DR   EMBL; M11522; AAB59551.1; JOINED.
DR   EMBL; M11523; AAB59551.1; JOINED.
DR   EMBL; M11524; AAB59551.1; JOINED.
DR   EMBL; M11525; AAB59551.1; JOINED.
DR   EMBL; M11526; AAB59551.1; JOINED.
DR   EMBL; M11527; AAB59551.1; JOINED.
DR   EMBL; M11528; AAB59551.1; JOINED.
DR   PIR; A01279; KGHUH1.
DR   PIR; A01280; KGHUL1.
DR   SWISS-2DPAGE; P01042; HUMAN.
DR   Genew; HGNC:6383; KNG.
DR   MIM; 228960; -.
DR   GO; GO:0007596; P:blood coagulation; NAS.
DR   GO; GO:0030146; P:diuresis; NAS.
DR   GO; GO:0006954; P:inflammatory response; NAS.
DR   GO; GO:0030147; P:natriuresis; NAS.
DR   GO; GO:0006939; P:smooth muscle contraction; NAS.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Pyrrolidone carboxylic acid.
FT   SIGNAL       1    18
FT   CHAIN       19   644       KININOGEN.
FT   CHAIN       19   380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381   389       BRADYKININ.
FT   CHAIN      390   644       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19   136       CYSTATIN-LIKE 1.
FT   DOMAIN     137   258       CYSTATIN-LIKE 2.
FT   DOMAIN     259   380       CYSTATIN-LIKE 3.
FT   DOMAIN     420   510       HIS-RICH
FT                              (ASSOCIATED WITH CLOTTING ACTIVITY).
FT   REPEAT     420   449
FT   REPEAT     450   479
FT   REPEAT     480   510
FT   MOD_RES     19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   DISULFID    28   614       INTERCHAIN.
FT   DISULFID    83    94
FT   DISULFID   107   126
FT   DISULFID   142   145
FT   DISULFID   206   218
FT   DISULFID   229   248
FT   DISULFID   264   267
FT   DISULFID   328   340
FT   DISULFID   351   370
FT   CARBOHYD    48    48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294   294       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   401   401       O-LINKED.
FT   CARBOHYD   533   533       O-LINKED.
FT   CARBOHYD   542   542       O-LINKED.
FT   CARBOHYD   546   546       O-LINKED.
FT   CARBOHYD   557   557       O-LINKED.
FT   CARBOHYD   571   571       O-LINKED.
FT   CARBOHYD   577   577       O-LINKED.
FT   CARBOHYD   593   593       O-LINKED.
FT   CARBOHYD   628   628       O-LINKED.
FT   VARSPLIC   402   427       VSPPHTSMAPAQDEERDSGKEQGHTR -> SHLRSCEYKGR
FT                              PPKAGAEPASEREVS (in isoform LMW).
FT                              /FTId=VSP_001261.
FT   VARSPLIC   428   644       Missing (in isoform LMW).
FT                              /FTId=VSP_001262.
FT   CONFLICT   593   593       T -> I (IN REF. 1).
SQ   SEQUENCE   644 AA;  71945 MW;  3132B4CBAF8FBB7E CRC64;

  Query Match           100.0%;  Score 660;  DB 1;  Length 644;
  Best Local Similarity  100.0%;  Pred. No. 2e-55;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 2
```

```
KNL1_BOVIN
ID   KNL1_BOVIN     STANDARD;      PRT;   436 AA.
AC   P01046;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen, LMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (3) the
CC       active peptide kallidin that is released from LMW-kininogen shows
CC       a variety of physiological effects: (3A) influence in smooth
CC       muscle contraction, (3B) induction of hypotension, (3C)
CC       natriuresis and diuresis (kidney).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=LMW I;
CC         IsoId=P01046-1; Sequence=Displayed;
CC       Name=HMW I;
CC         IsoId=P01044-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: LMW-kininogen is in contrast to HMW-kininogen not
CC       involved in blood clotting.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00426; CAA23709.1; -.
DR   PIR; A01283; KGBOL1.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     18
FT   CHAIN        19    436       KININOGEN, LMW I.
FT   CHAIN        19    378       HEAVY CHAIN.
FT   PEPTIDE     380    388       BRADYKININ.
FT   CHAIN       389    436       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    378       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    406       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    263    266
FT   DISULFID    327    339
FT   DISULFID    350    369
FT   CONFLICT    295    295       A -> T (IN REF. 1; CAA23709).
SQ   SEQUENCE   436 AA;  48427 MW;  F01F7EB6814BCE6C CRC64;

  Query Match             72.3%;  Score 477;  DB 1;  Length 436;
  Best Local Similarity   71.9%;  Pred. No. 3.6e-38;
  Matches   87;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              ||||||||::| |||: || :||:||| |:|:| ||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         62 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPL 121
              | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:|||||||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQPL 372

Qy        122 G 122
              |
Db        373 G 373


RESULT 3
KNH1_BOVIN
ID   KNH1_BOVIN     STANDARD;      PRT;   621 AA.
AC   P01044;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Kininogen, HMW I precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-378.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 378-393.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
```

```
RP   SEQUENCE OF 458-498.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW I;
CC         IsoId=P01044-1; Sequence=Displayed;
CC       Name=LMW I;
CC         IsoId=P01046-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01491; CAA24735.1; -.
DR   PIR; A01281; KGBOH1.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation;
KW   Inflammatory response; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL       1    18       PROBABLE.
FT   CHAIN       19   621       KININOGEN, HMW I.
FT   CHAIN       19   378       HEAVY CHAIN.
FT   PEPTIDE    380   388       BRADYKININ.
FT   CHAIN      389   621       LIGHT CHAIN.
FT   DOMAIN      19   135       CYSTATIN-LIKE 1.
FT   DOMAIN     136   257       CYSTATIN-LIKE 2.
FT   DOMAIN     258   378       CYSTATIN-LIKE 3.
FT   MOD_RES     19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    87    87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   136   136       O-LINKED (PARTIAL).
FT   CARBOHYD   168   168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD   197   197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   204   204       N-LINKED (GLCNAC. . .).
FT   DISULFID    27   591       INTERCHAIN.
FT   DISULFID    82    93
FT   DISULFID   106   125
FT   DISULFID   141   144
FT   DISULFID   205   217
FT   DISULFID   228   247
FT   DISULFID   263   266
FT   DISULFID   327   339
FT   DISULFID   350   369
SQ   SEQUENCE   621 AA;  68890 MW;  D16850BEFE3C55CD CRC64;

  Query Match             72.3%;  Score 477;  DB 1;  Length 621;
  Best Local Similarity   71.9%;  Pred. No. 5.4e-38;
  Matches   87;  Conservative   14;  Mismatches   20;  Indels    0;  Gaps    0;

Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              ||||||||::| |||: || :||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         62 GKKYPIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPL 121
              | || | |:|||||||| ||||||:|||     || | |:| |||||||:|:|||||||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQPL 372

Qy        122 G 122
              |
Db        373 G 373


RESULT 4
KNL2_BOVIN
ID   KNL2_BOVIN     STANDARD;      PRT;   434 AA.
AC   P01047;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83117859; PubMed=6572010;
RA   Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT   "Primary structures of bovine liver low molecular weight kininogen
RT   precursors and their two mRNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (3) the
CC       active peptide kallidin that is released from LMW-kininogen shows
CC       a variety of physiological effects: (3A) influence in smooth
CC       muscle contraction, (3B) induction of hypotension, (3C)
CC       natriuresis and diuresis (kidney).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=LMW II;
CC         IsoId=P01047-1; Sequence=Displayed;
CC       Name=HMW II;
CC         IsoId=P01045-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: LMW-kininogen is in contrast to HMW-kininogen not
CC       involved in blood clotting.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V00427; CAA23710.1; -.
DR   PIR; A01284; KGBOL2.
DR   HSSP; P01038; 1A90.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Signal;
KW   Pyrrolidone carboxylic acid.
FT   SIGNAL        1     18
FT   CHAIN        19    434       KININOGEN, LMW II.
FT   CHAIN        19    376       HEAVY CHAIN.
FT   PEPTIDE     378    386       BRADYKININ.
FT   CHAIN       387    434       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    256       CYSTATIN-LIKE 2.
FT   DOMAIN      257    376       CYSTATIN-LIKE 3.
FT   MOD_RES      19     19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    136    136       O-LINKED (PARTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .).
FT   DISULFID     27    404       INTERCHAIN.
FT   DISULFID     82     93
FT   DISULFID    106    125
FT   DISULFID    141    144
FT   DISULFID    205    217
FT   DISULFID    228    247
FT   DISULFID    261    264
FT   DISULFID    325    337
FT   DISULFID    348    367
SQ   SEQUENCE   434 AA;  48148 MW;  73A7079DE3E03430 CRC64;

  Query Match           68.2%;  Score 450;  DB 1;  Length 434;
  Best Local Similarity  68.9%;  Pred. No. 1.4e-35;
  Matches   84;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |:||:  |  :|||||: || :||:||| | |:| |||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
               | || | |:||||||| ||||||:|||    || | |:| |||||||:|:||||||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQP 369

Qy        121 LG 122
              ||
Db        370 LG 371


RESULT 5
KNH2_BOVIN
ID   KNH2_BOVIN     STANDARD;      PRT;   619 AA.
AC   P01045;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Kininogen, HMW II precursor (Thiol proteinase inhibitor) [Contains:
DE   Bradykinin].
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84014106; PubMed=6571699;
RA   Kitamura N., Takagaki Y., Furuto S., Tanaka T., Nawa H., Nakanishi S.;
RT   "A single gene for bovine high molecular weight and low molecular
RT   weight kininogens.";
RL   Nature 305:545-549(1983).
RN   [2]
RP   SEQUENCE OF 19-376.
RX   MEDLINE=87137530; PubMed=3546295;
RA   Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA   Miyata T., Iwanaga S.;
RT   "Bovine high molecular weight kininogen. The amino acid sequence,
RT   positions of carbohydrate chains and disulfide bridges in the heavy
RT   chain portion.";
RL   J. Biol. Chem. 262:2768-2779(1987).
RN   [3]
RP   SEQUENCE OF 376-391.
RX   MEDLINE=70180420; PubMed=4986212;
RA   Kato H., Nagasawa S., Suzuki T.;
RT   "Studies on the structure of bovine kininogen: cleavages of disulfide
RT   bonds and of methionyl bonds in kininogen-II.";
RL   J. Biochem. 67:313-323(1970).
RN   [4]
RP   SEQUENCE OF 387-455.
RX   MEDLINE=76260155; PubMed=956151;
RA   Han Y.N., Kato H., Iwanaga S., Suzuki T.;
RT   "Primary structure of bovine plasma high-molecular-weight kininogen.
RT   The amino acid sequence of a glycopeptide portion (fragment 1)
RT   following the C-terminus ot the bradykinin moiety.";
RL   J. Biochem. 79:1201-1222(1976).
RN   [5]
RP   SEQUENCE OF 456-496.
RX   MEDLINE=75170265; PubMed=1169237;
RA   Han Y.N., Komiya M., Iwanaga S., Suzuki T.;
RT   "Studies on the primary structure of bovine high-molecular-weight
RT   kininogen. Amino acid sequence of a fragment ('histidine-rich
RT   peptide') released by plasma kallikrein.";
RL   J. Biochem. 77:55-68(1975).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin- and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action).
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW II;
CC         IsoId=P01045-1; Sequence=Displayed;
CC       Name=LMW II;
CC         IsoId=P01047-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; V01492; CAA24736.1; -.
DR   PIR; A01282; KGBOH2.
DR   HSSP; P01038; 1A90.
```

```
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW   Thiol protease inhibitor; Bradykinin; Blood coagulation; Signal;
KW   Inflammatory response; Pyrrolidone carboxylic acid.
FT   SIGNAL       1    18
FT   CHAIN       19   619       KININOGEN, HMW II.
FT   CHAIN       19   376       HEAVY CHAIN.
FT   PEPTIDE    378   386       BRADYKININ.
FT   CHAIN      387   619       LIGHT CHAIN.
FT   DOMAIN      19   135       CYSTATIN-LIKE 1.
FT   DOMAIN     136   256       CYSTATIN-LIKE 2.
FT   DOMAIN     257   376       CYSTATIN-LIKE 3.
FT   MOD_RES     19    19       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    87    87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   136   136       O-LINKED (PARTIAL).
FT   CARBOHYD   168   168       N-LINKED (GLCNAC. . .) (OR 169).
FT   CARBOHYD   197   197       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   204   204       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   280   280       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   400   400       O-LINKED.
FT   DISULFID    27   589       INTERCHAIN.
FT   DISULFID    82    93
FT   DISULFID   106   125
FT   DISULFID   141   144
FT   DISULFID   205   217
FT   DISULFID   228   247
FT   DISULFID   261   264
FT   DISULFID   325   337
FT   DISULFID   348   367
FT   VARIANT    398   398       T -> P.
FT   VARIANT    401   401       L -> V.
FT   VARIANT    454   454       H -> K.
SQ   SEQUENCE   619 AA;  68710 MW;  F04320A8EB0EE0DA CRC64;

  Query Match             68.2%;  Score 450;  DB 1;  Length 619;
  Best Local Similarity   68.9%;  Pred. No. 2e-35;
  Matches   84;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |:||:  |  :|||||: || :||:||| | |:| ||||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:||||||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQP 369

Qy        121 LG 122
              ||
Db        370 LG 371


RESULT 6
KNG_MOUSE
ID   KNG_MOUSE      STANDARD;      PRT;   661 AA.
AC   O08677; O08676; Q91XK5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC   STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX   MEDLINE=97342556; PubMed=9199253;
RA   Takano M., Kondo J., Yayama K., Otani M., Sano K., Okamoto H.;
RT   "Molecular cloning of cDNAs for mouse low-molecular-weight and high-
RT   molecular-weight prekininogens.";
RL   Biochim. Biophys. Acta 1352:222-230(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RC   STRAIN=C57BL/6J; TISSUE=Placenta;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RC   TISSUE=Liver;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
```

```
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
CC       clotting (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=O08677-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=O08677-2; Sequence=VSP_001263, VSP_001264;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84435; BAA19743.1; -.
DR   EMBL; D84415; BAA19742.1; -.
DR   EMBL; AK005547; BAB24115.1; -.
DR   EMBL; BC018158; AAH18158.1; -.
DR   MGD; MGI:1097705; Kng.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    661       KININOGEN.
FT   CHAIN        19    379       KININOGEN HEAVY CHAIN.
FT   CHAIN       380    388       BRADYKININ.
FT   CHAIN       389    661       KININOGEN LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    379       CYSTATIN-LIKE 3.
FT   DOMAIN      439    524       HIS-RICH.
FT   DISULFID     28    631       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    401    432       VSPPYIAREQEERDAETEQGPTHGHGWLHEKQ -> RLLRA
FT                                CEYKGRLSKAGAEPAPERQAESSQVKQ (in isoform
FT                                LMW).
FT                                /FTId=VSP_001263.
FT   VARSPLIC    433    661       Missing (in isoform LMW).
FT                                /FTId=VSP_001264.
SQ   SEQUENCE   661 AA;  73102 MW;  774460258D58796E CRC64;

  Query Match             65.2%;  Score 430;  DB 1;  Length 661;
  Best Local Similarity   65.9%;  Pred. No. 1.8e-33;
  Matches   81;  Conservative   11;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | | |:   | | ||||||| :||||:| | |:| :|||||:  ||:||| ||||  |||
Db        252 GDDLVEALPKPCPGCPRDIPVDSPELKEVLGHSIAQLNAENDHPFYYKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||||| || | || | |||||||||| ||: ||| |: ||| ||
Db        312 AGTKYVIEFIARETKCSKESNTELAEDCEIKHLGQSLDCNANVYMRPWENKVVPTVKCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374


RESULT 7
KNG_RAT
ID   KNG_RAT        STANDARD;      PRT;   639 AA.
AC   P08934; P08933;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Kininogen precursor [Contains: Bradykinin].
GN   KNG.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RX   MEDLINE=87137443; PubMed=3029068;
RA   Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
RT   "Differing expression patterns and evolution of the rat kininogen
RT   gene family.";
RL   J. Biol. Chem. 262:2190-2198(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM LMW).
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [3]
RP   SEQUENCE OF 1-65 FROM N.A.
RC   STRAIN=Buffalo;
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
RN   [4]
RP   SEQUENCE OF 1-41 FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=87137465; PubMed=3818598;
RA   Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT   "Differing utilization of homologous transcription initiation sites
RT   of rat K and T kininogen genes under inflammation condition.";
RL   J. Biol. Chem. 262:2345-2351(1987).
CC   -!- FUNCTION: (1) Kininogens are inhibitors of thiol proteases; (2)
CC       HMW-kininogen plays an important role in blood coagulation by
CC       helping to position optimally prekallikrein and factor XI next to
CC       factor XII; (3) HMW-kininogen inhibits the thrombin-and plasmin-
CC       induced aggregation of thrombocytes; (4) the active peptide
CC       bradykinin that is released from HMW-kininogen shows a variety of
CC       physiological effects: (4A) influence in smooth muscle
CC       contraction, (4B) induction of hypotension, (4C) natriuresis and
CC       diuresis, (4D) decrease in blood glucose level, (4E) it is a
CC       mediator of inflammation and causes (4E1) increase in vascular
CC       permeability, (4E2) stimulation of nociceptors (4E3) release of
CC       other mediators of inflammation (e.g. prostaglandins), (4F) it has
CC       a cardioprotective effect (directly via bradykinin action,
CC       indirectly via endothelium-derived relaxing factor action); (5)
CC       LMW-kininogen inhibits the aggregation of thrombocytes; (6) LMW-
CC       kininogen is in contrast to HMW-kininogen not involved in blood
```

```
CC       clotting.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=HMW;
CC         IsoId=P08934-1; Sequence=Displayed;
CC       Name=LMW;
CC         IsoId=P08934-2; Sequence=VSP_001265, VSP_001266;
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Bradykinin is released from kininogen by plasma kallikrein.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW/LMW kininogens and two additional LMW-like
CC       kininogens: T-I and T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29428; AAA41486.1; -.
DR   EMBL; M11884; AAA41487.1; -.
DR   EMBL; M14369; AAA41484.1; -.
DR   EMBL; M14369; AAA41485.1; ALT_SEQ.
DR   EMBL; M16455; AAA41482.1; -.
DR   PIR; A25486; A25486.
DR   PIR; A28055; A28055.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR002395; Kininogen.
DR   Pfam; PF00031; cystatin; 3.
DR   PRINTS; PR00334; KININOGEN.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW   Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW   Alternative splicing; Multigene family.
FT   SIGNAL       1    18
FT   CHAIN       19   639       KININOGEN.
FT   CHAIN       19   380       KININOGEN HEAVY CHAIN.
FT   PEPTIDE    381   389       BRADYKININ.
FT   CHAIN      390   639       KININOGEN LIGHT CHAIN.
FT   DOMAIN      19   136       CYSTATIN-LIKE 1.
FT   DOMAIN     137   258       CYSTATIN-LIKE 2.
FT   DOMAIN     259   380       CYSTATIN-LIKE 3.
FT   DOMAIN     439   514       HIS-RICH.
FT   DISULFID    28   609       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    83    94       BY SIMILARITY.
FT   DISULFID   107   126       BY SIMILARITY.
FT   DISULFID   142   145       BY SIMILARITY.
FT   DISULFID   206   218       BY SIMILARITY.
FT   DISULFID   229   248       BY SIMILARITY.
FT   DISULFID   264   267       BY SIMILARITY.
FT   DISULFID   328   340       BY SIMILARITY.
FT   DISULFID   351   370       BY SIMILARITY.
FT   CARBOHYD    82    82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   127   127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   169   169       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   205   205       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294   294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   529   529       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   402   433       VSPSYIARVQEERDPGNEQGPIHGHGWLHAKQ -> RLLNS
FT                              CEYKGRLLKAGAGPAPERQAEASTVTP (in isoform
FT                              LMW).
FT                              /FTId=VSP_001265.
FT   VARSPLIC   434   639       Missing (in isoform LMW).
FT                              /FTId=VSP_001266.
FT   CONFLICT    61    61       E -> K (IN REF. 2).
SQ   SEQUENCE   639 AA;  70933 MW;  D3172DF94FF56AF5 CRC64;

  Query Match             64.5%;  Score 426;  DB 1;  Length 639;
  Best Local Similarity   65.0%;  Pred. No. 4.1e-33;
  Matches   80;  Conservative   14;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |  |   :   : | ||||:|| :||||:| | |:| :|||||| ||||||| ||||  |||
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||||| ||| ||||:|||||:||| ||: ||| |: ||| |:
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTVKCKV 372

Qy        121 LGM 123
              |.|
Db        373 LDM 375


RESULT 8
KNT2_RAT
ID   KNT2_RAT       STANDARD;      PRT;   430 AA.
AC   P08932;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   T-kininogen II precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
CC   -!- FUNCTION: Kininogens are plasma glycoproteins with a number of
CC       functions: (1) as precursor of the active peptide bradykinin they
CC       effect smooth muscle contraction, induction of hypotension and
CC       increase of vascular permeability. (2) They play a role in blood
CC       coagulation by helping to position optimally prekallikrein and
CC       factor XI next to factor XII. (3) They are inhibitor of thiol
CC       proteases.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- INDUCTION: In response to an inflammatory stimulant. T-kininogen
CC       II synthesis is induced and the plasma concentration of
CC       T-kininogen I is raised.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW and LMW kininogens produced by alternative splicing
CC       of the same gene, and two additional LMW-like kininogens: T-I and
CC       T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11885; AAA41491.1; -.
DR   PIR; B28055; B28055.
DR   GlycoSuiteDB; P08932; -.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
```

```
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-II.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   430 AA;  47524 MW;  43EDF02D1BF55076 CRC64;

  Query Match             62.0%;  Score 409;  DB 1;  Length 430;
  Best Local Similarity   61.8%;  Pred. No. 1.1e-31;
  Matches   76;  Conservative   15;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |       | | |||::|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||::| |||  |||| |||||:||| ||: ||| |: ||| ||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374


RESULT 9
KNT1_RAT
ID   KNT1_RAT       STANDARD;      PRT;   430 AA.
AC   P01048; P04081;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   T-kininogen I precursor (Major acute phase protein) (Alpha-1-MAP)
DE   (Thiostatin) [Contains: T-kinin].
GN   MAP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86008264; PubMed=2413018;
RA   Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT   "Primary structures of the mRNAs encoding the rat precursors for
RT   bradykinin and T-kinin. Structural relationship of kininogens with
RT   major acute phase protein and alpha 1-cysteine proteinase
RT   inhibitor.";
RL   J. Biol. Chem. 260:12054-12059(1985).
RN   [2]
RP   SEQUENCE OF 5-430 FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=86008266; PubMed=2413019;
RA   Anderson K.P., Heath E.C.;
RT   "The relationship between rat major acute phase protein and the
RT   kininogens.";
RL   J. Biol. Chem. 260:12065-12071(1985).
RN   [3]
RP   SEQUENCE OF 7-430 FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha 1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
RN   [4]
RP   SEQUENCE OF 1-65 FROM N.A.
RX   MEDLINE=87250580; PubMed=2439509;
RA   Fung W.-P., Schreiber G.;
RT   "Structure and expression of the genes for major acute phase alpha 1-
RT   protein (thiostatin) and kininogen in the rat.";
RL   J. Biol. Chem. 262:9298-9308(1987).
CC   -!- FUNCTION: Kininogens are plasma glycoproteins with a number of
CC       functions: (1) as precursor of the active peptide bradykinin they
CC       effect smooth muscle contraction, induction of hypotension and
CC       increase of vascular permeability. (2) They play a role in blood
CC       coagulation by helping to position optimally prekallikrein and
CC       factor XI next to factor XII. (3) They are inhibitor of thiol
CC       proteases.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- INDUCTION: In response to an inflammatory stimulant. T-kininogen
CC       II synthesis is induced and the plasma concentration of
CC       T-kininogen I is raised.
CC   -!- PTM: AS T-KININ IS PRECEDED BY A MET INSTEAD OF AN ARG OR LYS, IT
CC       IS NOT RELEASED FROM ITS PRECURSOR BY EITHER TISSUE OR PLASMA
CC       KALLIKREIN.
CC   -!- MISCELLANEOUS: Rats express four types of kininogens: the
CC       classical HMW and LMW kininogens produced by alternative splicing
CC       of the same gene, and two additional LMW-like kininogens: T-I and
CC       T-II.
CC   -!- SIMILARITY: Contains 3 cystatin-like domains.
CC   -!- CAUTION: In addition to the conflicts described in the feature
CC       table, Ref.2 sequence differs from that shown in positions 257,
CC       262, 268, 269, 295, 314, 315, 331, 332 and 389. In all those
CC       positions the alternate amino acid is the one present in T-II
CC       kininogen.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11883; AAA41489.1; -.
DR   EMBL; M11661; AAA41570.1; -.
DR   EMBL; M16454; AAA41568.1; -.
DR   EMBL; X02299; CAA26162.1; ALT_SEQ.
DR   PIR; A01286; KGRTT1.
DR   PIR; A23897; A23897.
DR   PIR; A27115; A27115.
DR   GlycoSuiteDB; P01048; -.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Glycoprotein; Plasma; Repeat; Vasodilator; Multigene family;
KW   Thiol protease inhibitor; Bradykinin; Acute phase; Signal.
FT   SIGNAL        1     18
FT   CHAIN        19    430       KININOGEN, T-I.
FT   CHAIN        19    375       HEAVY CHAIN.
FT   PEPTIDE     376    386       T-KININ.
FT   CHAIN       387    430       LIGHT CHAIN.
FT   DOMAIN       19    135       CYSTATIN-LIKE 1.
FT   DOMAIN      136    257       CYSTATIN-LIKE 2.
FT   DOMAIN      258    375       CYSTATIN-LIKE 3.
FT   DISULFID     28    404       INTERCHAIN (BY SIMILARITY).
```

```
FT   DISULFID     83     94       BY SIMILARITY.
FT   DISULFID    107    125       BY SIMILARITY.
FT   DISULFID    141    144       BY SIMILARITY.
FT   DISULFID    205    217       BY SIMILARITY.
FT   DISULFID    228    247       BY SIMILARITY.
FT   DISULFID    263    266       BY SIMILARITY.
FT   DISULFID    327    339       BY SIMILARITY.
FT   DISULFID    350    369       BY SIMILARITY.
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    168    168       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     26     28       LNC -> MDR (IN REF. 2).
FT   CONFLICT     55     55       V -> L (IN REF. 2).
FT   CONFLICT     61     61       E -> K (IN REF. 1).
FT   CONFLICT     83     83       C -> Y (IN REF. 3).
FT   CONFLICT    166    166       S -> F (IN REF. 2 AND 3).
FT   CONFLICT    179    181       REV -> TKI (IN REF. 2).
FT   CONFLICT    193    193       N -> D (IN REF. 2).
FT   CONFLICT    212    212       S -> F (IN REF. 2).
FT   CONFLICT    214    214       R -> H (IN REF. 3).
FT   CONFLICT    229    229       T -> R (IN REF. 2).
FT   CONFLICT    233    233       H -> Y (IN REF. 2).
FT   CONFLICT    257    257       E -> S (IN REF. 2).
FT   CONFLICT    262    262       N -> K (IN REF. 2).
FT   CONFLICT    264    264       R -> F (IN REF. 2).
FT   CONFLICT    268    269       RE -> KN (IN REF. 2).
FT   CONFLICT    295    295       I -> L (IN REF. 2).
FT   CONFLICT    314    315       VI -> TK (IN REF. 2).
FT   CONFLICT    331    332       SK -> TN (IN REF. 2).
FT   CONFLICT    389    389       R -> Q (IN REF. 2).
FT   CONFLICT    414    414       R -> G (IN REF. 2 AND 3).
FT   CONFLICT    415    415       A -> L (IN REF. 2).
FT   CONFLICT    420    421       DH -> ER (IN REF. 3).
FT   CONFLICT    430    430       P -> S (IN REF. 1).
SQ   SEQUENCE   430 AA;  47715 MW;  FAEBB78FAF4723C3 CRC64;

  Query Match             60.8%;  Score 401;  DB 1;  Length 430;
  Best Local Similarity   61.8%;  Pred. No. 6.3e-31;
  Matches   76;  Conservative   14;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: ||| ||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374


RESULT 10
CYTF_MOUSE
ID   CYTF_MOUSE     STANDARD;      PRT;   144 AA.
AC   O89098;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin F precursor (Leukocystatin) (Cystatin 7) (Cystatin-like
DE   metastasis-associated protein) (CMAP).
GN   CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98298157; PubMed=9632704;
RA   Halfon S., Ford J., Foster J., Dowling L., Lucian L., Sterling M.,
RA   Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,
RA   Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.;
RT   "Leukocystatin, a new class II cystatin expressed selectively by
RT   hematopoietic cells.";
RL   J. Biol. Chem. 273:16400-16408(1998).
CC   -!- FUNCTION: Inhibits papain and cathepsin L but with affinities
CC       lower than other cystatins. May play a role in immune regulation
CC       through inhibition of a unique target in the hematopoietic system.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF031826; AAC40140.1; -.
DR   EMBL; AF031825; AAC40139.1; -.
DR   HSSP; P01034; 1G96.
DR   MGD; MGI:1298217; Cst7.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW   Thiol protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    144       CYSTATIN F.
FT   ACT_SITE    36     36       REACTIVE SITE.
FT   SITE        80     84       SECONDARY AREA OF CONTACT.
FT   DISULFID    98    109       BY SIMILARITY.
FT   DISULFID   123    143       BY SIMILARITY.
SQ   SEQUENCE   144 AA;  16380 MW;  B5837334C1B4A89C CRC64;

  Query Match             26.5%;  Score 175;  DB 1;  Length 144;
  Best Local Similarity   34.4%;  Pred. No. 6.8e-10;
  Matches   42;  Conservative   23;  Mismatches   49;  Indels    8;  Gaps    4;

Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              ||  :   : :  | |: | ||:| : :   |:: | |   |  | ||  :| || ||||
Db         27 KDLI---SSVKPGFPKTIETNNPGVLKAARHSVEKFNNCTNDIFLFKESHVSKALVQVVK 83

Qy         62 GKKYFIDFVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVN 117
              | || ::     ||| |  : :| ::|:     | ::| | :||:|:||     | :
Db         84 GLKYMLEVKIGRTTCRKTMHHQL-DNCDFQTNPALKRTLYCYSEVWVIPWLHSFEVPVLL 142

Qy        118 CQ 119
              ||
Db        143 CQ 144


RESULT 11
CYTF_HUMAN
ID   CYTF_HUMAN     STANDARD;      PRT;   145 AA.
AC   O76096; Q9UED4;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cystatin F precursor (Leukocystatin) (Cystatin 7) (Cystatin-like
DE   metastasis-associated protein) (CMAP).
GN   CST7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98406133; PubMed=9733783;
RA   Ni J., Fernandez M.A., Danielsson L., Chillakuru R.A., Zhang J.,
RA   Grubb A., Su J., Gentz R., Abrahamson M.;
RT   "Cystatin F is a glycosylated human low molecular weight cysteine
```

```
RT   proteinase inhibitor.";
RL   J. Biol. Chem. 273:24797-24804(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98298157; PubMed=9632704;
RA   Halfon S., Ford J., Foster J., Dowling L., Lucian L., Sterling M.,
RA   Xu Y., Weiss M., Ikeda M., Liggett D., Helms A., Caux C., Lebecque S.,
RA   Hannum C., Menon S., McClanahan T., Gorman D., Zurawski G.;
RT   "Leukocystatin, a new class II cystatin expressed selectively by
RT   hematopoietic cells.";
RL   J. Biol. Chem. 273:16400-16408(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Morita M., Arakawa H., Yoshiuchi N.;
RT   "Human homologue of murine CMAP.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20399571; PubMed=10945474;
RA   Morita M., Hara Y., Tamai Y., Arakawa H., Nishimura S.;
RT   "Genomic construct and mapping of the gene for CMAP
RT   (Leukocystatin/Cystatin F, CST7) and identification of a proximal
RT   novel gene, BSCv (C20orf3).";
RL   Genomics 67:87-91(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Inhibits papain and cathepsin L but with affinities
CC       lower than other cystatins. May play a role in immune regulation
CC       through inhibition of a unique target in the hematopoietic system.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- TISSUE SPECIFICITY: Primarily expressed in peripheral blood cells
CC       and spleen.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF036342; AAC35747.1; -.
DR   EMBL; AF031824; AAC39788.1; -.
DR   EMBL; AB015225; BAA34941.1; ALT_INIT.
DR   EMBL; AB029636; BAB11886.1; ALT_INIT.
DR   EMBL; AL035661; CAB75498.1; -.
DR   EMBL; BC015507; AAH15507.1; ALT_INIT.
DR   HSSP; P01034; 1G96.
DR   Genew; HGNC:2479; CST7.
DR   MIM; 603253; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; FALSE_NEG.
KW   Thiol protease inhibitor; Glycoprotein; Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    145       CYSTATIN F.
FT   ACT_SITE     37     37       REACTIVE SITE.
FT   SITE         81     85       SECONDARY AREA OF CONTACT.
FT   DISULFID     99    110       BY SIMILARITY.
FT   DISULFID    124    144       BY SIMILARITY.
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   145 AA;  16454 MW;  B2BCC4F76857CB0F CRC64;

  Query Match             25.1%;  Score 165.5;  DB 1;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 5.5e-09;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |   | | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                 ||| |  :  | : |:      | :| | :||:|||| :    | : |
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 12
CYTM_HUMAN
ID   CYTM_HUMAN     STANDARD;      PRT;   149 AA.
AC   Q15828;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Cystatin M precursor (Cystatin E).
GN   CST6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97150844; PubMed=8995380;
RA   Sotiropoulou G., Anisowicz A., Sager R.;
RT   "Identification, cloning, and characterization of cystatin M, a novel
RT   cysteine proteinase inhibitor, down-regulated in breast cancer.";
RL   J. Biol. Chem. 272:903-910(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97256812; PubMed=9099741;
RA   Ni J., Abrahamson M., Zhang M., Fernandez M.A., Grubb A., Su J.,
RA   Yu G.L., Li Y., Parmelee D., Xing L., Coleman T.A., Gentz S.,
RA   Thotakura R., Nguyen N., Hesselberg M., Gentz R.;
RT   "Cystatin E is a novel human cysteine proteinase inhibitor with
RT   structural resemblance to family 2 cystatins.";
RL   J. Biol. Chem. 272:10853-10858(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   CHARACTERIZATION, AND TISSUE SPECIFICITY.
RX   MEDLINE=21246880; PubMed=11348457;
RA   Zeeuwen P.L., Van Vlijmen-Willems I.M., Jansen B.J., Sotiropoulou G.,
RA   Curfs J.H., Meis J.F., Janssen J.J., Van Ruissen F., Schalkwijk J.;
RT   "Cystatin M/E expression is restricted to differentiated epidermal
RT   keratinocytes and sweat glands: a new skin-specific proteinase
RT   inhibitor that is a target for cross-linking by transglutaminase.";
RL   J. Invest. Dermatol. 116:693-701(2001).
CC   -!- FUNCTION: Shows moderate inhibition of cathepsin B but is not
CC       active against cathepsin C.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Restricted to the stratum granulosum of normal
CC       skin, the stratum granulosum/spinosum of psoriatic skin, and the
CC       secretory coils of eccrine sweat glands. Low expression levels are
CC       found in the nasal cavity.
CC   -!- PTM: Substrate for transglutaminases. Acts as an acyl acceptor but
CC       not as an acyl donor.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U62800; AAB06566.1; -.
DR   EMBL; U81233; AAB61305.1; -.
DR   EMBL; BC031334; AAH31334.1; -.
DR   HSSP; P01038; 1CEW.
DR   Genew; HGNC:2478; CST6.
DR   MIM; 601891; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Signal; Glycoprotein.
FT   SIGNAL        1     28       PROBABLE.
FT   CHAIN        29    149       CYSTATIN M.
FT   ACT_SITE     36     36       REACTIVE SITE.
FT   SITE         80     84       SECONDARY AREA OF CONTACT.
FT   DISULFID     98    113       BY SIMILARITY.
FT   DISULFID    126    146       BY SIMILARITY.
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   149 AA;  16511 MW;  2076A78BFC9FAC8C CRC64;

  Query Match             21.0%;  Score 138.5;  DB 1;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 2.1e-06;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
              |  :  ||  ||:  :  |:::::     :    |  :|: :||:  :: ||: |:||| |||:
Db         30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy         68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                 | | |      : :|| :|      |   | |: || ||||:
Db         90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 13
CYTC_BOVIN
ID   CYTC_BOVIN     STANDARD;      PRT;   148 AA.
AC   P01035;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin C precursor (Colostrum thiol proteinase inhibitor).
GN   CST3.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 66-83, AND CHARACTERIZATION.
RC   TISSUE=Cerebrospinal fluid, and Choroid plexus;
RX   MEDLINE=98094199; PubMed=9434110;
RA   Olsson S.-L., Ek B., Wilm M., Broberg S., Rask L., Bjcerk I.;
RT   "Molecular cloning and N-terminal analysis of bovine cystatin C
RT   identification of a full-length N-terminal region.";
RL   Biochim. Biophys. Acta 1343:203-210(1997).
RN   [2]
RP   SEQUENCE OF 37-148.
RX   MEDLINE=85231205; PubMed=3891407;
RA   Hirado M., Tsunasawa S., Sakiyama F., Niinobe M., Fujii S.;
RT   "Complete amino acid sequence of bovine colostrum low-Mr cysteine
RT   proteinase inhibitor.";
RL   FEBS Lett. 186:41-45(1985).
CC   -!- FUNCTION: This is a thiol proteinase inhibitor.
CC   -!- MASS SPECTROMETRY: MW=13420; METHOD=MALDI.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y10811; CAA71771.1; -.
DR   HSSP; P01034; 1G96.
```

```
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Signal; Pyrrolidone carboxylic acid.
FT   SIGNAL        1     30       PROBABLE.
FT   CHAIN        31    148       CYSTATIN C.
FT   MOD_RES      31     31       PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
FT   ACT_SITE     40     40       REACTIVE SITE.
FT   SITE         84     88       SECONDARY AREA OF CONTACT.
FT   DISULFID    102    112       BY SIMILARITY.
FT   DISULFID    126    146       BY SIMILARITY.
SQ   SEQUENCE   148 AA;  16265 MW;  EE740FE37CFB9F0E CRC64;

  Query Match             20.6%;  Score 136;  DB 1;  Length 148;
  Best Local Similarity   30.6%;  Pred. No. 3.7e-06;
  Matches   34;  Conservative   22;  Mismatches   31;  Indels   24;  Gaps    5;

Qy         22 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 81
              |   ::| |:  ::: |  :|  :  ::  | :|| |||:|  ||:|     |||:|  :
Db         48 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 105

Qy         82 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVN-----CQ 119
              :   :||        : :||      |: :||||||      |:|     ||
Db        106 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 147


RESULT 14
CYTC_MACMU
ID   CYTC_MACMU     STANDARD;      PRT;   146 AA.
AC   O19092;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin C precursor.
GN   CST3.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97054523; PubMed=8898820;
RA   Wei L.H., Walker L.C., Levy E.;
RT   "Cystatin C. Icelandic-like mutation in an animal model of
RT   cerebrovascular beta-amyloidosis.";
RL   Stroke 27:2080-2085(1996).
CC   -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC       thought to serve an important physiological role as a local
CC       regulator of this enzyme activity.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U51912; AAB64050.1; -.
DR   HSSP; P01034; 1G96.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Amyloid; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    146       CYSTATIN C.
FT   ACT_SITE     37     37       REACTIVE SITE.
FT   SITE         81     85       SECONDARY AREA OF CONTACT.
FT   DISULFID     99    109       BY SIMILARITY.
FT   DISULFID    123    143       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  15857 MW;  F0B3BB774A29DF26 CRC64;

  Query Match             20.2%;  Score 133.5;  DB 1;  Length 146;
  Best Local Similarity   29.0%;  Pred. No. 6.3e-06;
  Matches   31;  Conservative   21;  Mismatches   48;  Indels    7;  Gaps    3;

Qy          6 QPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKY 65
              :||    || | |       :   |   ::: |  :|  :: :   | :|| |:|||  |
Db         31 KPPR--LVGGPMDASVEEEGVRRALDFAVSEYNKASNDMYHSRALQVVRARKQIVAGVNY 88

Qy         66 FIDFVARETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWE 109
              |:|      |||:|   :   ::|   |   | :   |: ::| |||:
Db         89 FLDVELGRTTCTK--TQPNLDNCPFHEQPHLKRKAFCSFQIYTVPWQ 133


RESULT 15
CYTC_SAISC
ID   CYTC_SAISC     STANDARD;      PRT;   146 AA.
AC   O19093;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cystatin C precursor.
GN   CST3.
OS   Saimiri sciureus (Common squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=9521;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97054523; PubMed=8898820;
RA   Wei L.H., Walker L.C., Levy E.;
RT   "Cystatin C. Icelandic-like mutation in an animal model of
RT   cerebrovascular beta-amyloidosis.";
RL   Stroke 27:2080-2085(1996).
CC   -!- FUNCTION: As an inhibitor of cysteine proteinases, this protein is
CC       thought to serve an important physiological role as a local
CC       regulator of this enzyme activity.
CC   -!- SIMILARITY: Belongs to the cystatin family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U52028; AAB64051.1; -.
DR   HSSP; P01034; 1G96.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
KW   Thiol protease inhibitor; Amyloid; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    146       CYSTATIN C.
FT   ACT_SITE     37     37       REACTIVE SITE.
FT   SITE         81     85       SECONDARY AREA OF CONTACT.
FT   DISULFID     99    109       BY SIMILARITY.
FT   DISULFID    123    143       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  15946 MW;  08196353C0306AA3 CRC64;

  Query Match             20.1%;  Score 132.5;  DB 1;  Length 146;
  Best Local Similarity   29.0%;  Pred. No. 7.8e-06;
  Matches   29;  Conservative   21;  Mismatches   45;  Indels    5;  Gaps    2;

Qy         13 VGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVAR 72
              :| | |       :   |   ::: |  :|  :: :   | :|| |:|||  ||:|
Db         36 LGGPMDASVEEEGVRRALDFAVSEYNKASNDMYHSRALQVVRARKQIVAGVNYFLDVEMG 95

Qy         73 ETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWE 109
```

```
                 |||:|  |:   ::|   |   | :   |: ::| |||:
Db          96 RTTCTK--NQPNLDNCPFHEQPHLKRKAFCSFQIYSVPWQ 133
```

Search completed: September 24, 2004, 14:09:12
Job time : 23.364 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:05:18 ; Search time 35.424 Seconds
(without alignments)
1095.549 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	404	61.2	140	6	Q7YRP6	Q7yrp6 sus scrofa
2	402	60.9	423	11	P70517	P70517 rattus norv
3	399	60.5	430	11	Q63581	Q63581 rattus norv
4	175	26.5	167	11	Q9QWL5	Q9qwl5 mus musculu
5	165.5	25.1	167	4	Q7Z4J8	Q7z4j8 homo sapien
6	152.5	23.1	462	13	Q7ZY91	Q7zy91 xenopus lae
7	152.5	23.1	462	13	Q7SYH2	Q7syh2 xenopus lae
8	152.5	23.1	465	13	Q801E5	Q801e5 xenopus lae
9	129.5	19.6	140	11	Q9EPX9	Q9epx9 mus musculu
10	122.5	18.6	455	13	Q800S8	Q800s8 brachydanio
11	118	17.9	464	13	Q801Z5	Q801z5 cyprinus ca
12	115.5	17.5	388	11	Q8CB17	Q8cb17 mus musculu
13	113.5	17.2	148	5	Q9NH95	Q9nh95 litomosoide
14	113	17.1	140	11	Q80Y72	Q80y72 mus musculu
15	111	16.8	146	11	Q8K397	Q8k397 mus musculu
16	111	16.8	149	11	Q9D1B1	Q9d1b1 mus musculu
17	108	16.4	112	13	Q98SR4	Q98sr4 acipenser s
18	108	16.4	112	13	Q98SR3	Q98sr3 acipenser s
19	107	16.2	139	5	Q9TYY2	Q9tyy2 caenorhabdi
20	106	16.1	300	13	Q801Z6	Q801z6 cyprinus ca
21	105	15.9	109	5	Q9TY65	Q9ty65 onchocerca
22	104.5	15.8	148	11	Q9JM84	Q9jm84 mus musculu
23	104.5	15.8	149	11	Q8VHC1	Q8vhc1 rattus norv
24	104.5	15.8	161	5	O16159	O16159 brugia mala
25	102.5	15.5	127	5	Q9U9A1	Q9u9a1 onchocerca
26	100	15.2	127	5	P90698	P90698 brugia mala
27	99	15.0	110	10	Q8SA65	Q8sa65 sandersonia
28	98.5	14.9	107	5	Q8T0Y2	Q8t0y2 sarcophaga
29	98.5	14.9	125	5	Q25620	Q25620 onchocerca
30	97.5	14.8	134	10	Q41825	Q41825 zea mays (m
31	95.5	14.5	143	5	O61973	O61973 caenorhabdi
32	94	14.2	122	5	O44396	O44396 haemonchus
33	93.5	14.2	138	4	Q8WXU6	Q8wxu6 homo sapien
34	90.5	13.7	157	5	Q17108	Q17108 acanthochei
35	90	13.6	139	11	Q8K5A3	Q8k5a3 rattus norv
36	88.5	13.4	92	10	Q9FXN6	Q9fxn6 arabidopsis
37	88.5	13.4	116	10	Q8RXS7	Q8rxs7 arabidopsis
38	88.5	13.4	124	10	Q41906	Q41906 arabidopsis
39	88.5	13.4	125	10	O22202	O22202 arabidopsis
40	88.5	13.4	134	10	P93627	P93627 zea mays (m
41	88.5	13.4	134	10	Q41897	Q41897 zea mays (m
42	87.5	13.3	141	11	Q9DAP1	Q9dap1 mus musculu
43	86.5	13.1	141	11	Q80ZN5	Q80zn5 mus musculu
44	85.5	13.0	144	5	Q966W0	Q966w0 nippostrong
45	85	12.9	199	10	Q39270	Q39270 brassica ca

ALIGNMENTS

RESULT 1
Q7YRP6
ID Q7YRP6 PRELIMINARY; PRT; 140 AA.
AC Q7YRP6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Low molecular weight kininogen (Fragment).
GN KNG.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Vonnahme K.A., Fernando S.C., Ross J.A., Ashworth M.D., DeSilva U.,
RA Malayer J.R., Geisert R.D.;
RT "Porcine Endometrial and Conceptus Expression of Kininogens and Plasma
RT Kallikrein in Cyclic and Pregnant Gilts.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY321363; AAP85260.1; -.
FT NON_TER 1 1
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15650 MW; 177837836603F777 CRC64;

Query Match 61.2%; Score 404; DB 6; Length 140;
Best Local Similarity 78.2%; Pred. No. 5.4e-33;
Matches 79; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

```
Qy         22 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 81
              :||:||| | |:| |||||||| ||||| |:|| ||||||||| | | |||||||:||
Db          1 DSPDLEEPLNHSIAKLNAENNAVFYFKIGPVEKATVQVVAGKKYSIVFTARETTCSKQSN 60

Qy         82 EELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLG 122
              |||||||| || || | ||| |||||||||||||| || ||
Db         61 EELTESCEIKKPGQILKCNASVYVVPWEKKIYPTVKCQLLG 101
```

RESULT 2

```
P70517
ID   P70517      PRELIMINARY;      PRT;   423 AA.
AC   P70517;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major acute phase alpha-1 protein precursor (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cole T.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85149311; PubMed=2579644;
RA   Cole T., Inglis A., Nagashima M., Schreiber G.;
RT   "Major acute-phase alpha(1)-protein in the rat: Structure, molecular
RT   cloning, and regulation of mRNA levels.";
RL   Biochem. Biophys. Res. Commun. 126:719-724(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85127561; PubMed=2578992;
RA   Cole T., Inglis A.S., Roxburgh C.M., Howlett G.J., Schreiber G.;
RT   "Major acute phase alpha1-protein of the rat is homologous to bovine
RT   kininogen and contains the sequence for bradykinin: its synthesis is
RT   regulated at the mRNA level.";
RL   FEBS Lett. 182:57-61(1985).
DR   EMBL; K02814; AAA41569.1; -.
DR   PIR; S68034; S68034.
DR   PIR; S68035; S68035.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     11       POTENTIAL.
FT   CHAIN         12    423       POTENTIAL.
FT   CHAIN        371    379       POTENTIAL.
SQ   SEQUENCE   423 AA;  46905 MW;  F9E8BD3198547949 CRC64;

  Query Match             60.9%;  Score 402;  DB 11;  Length 423;
  Best Local Similarity   61.8%;  Pred. No. 3.1e-32;
  Matches   76;  Conservative   14;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        245 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| ||||||:||| ||: ||| |: |||||
Db        305 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 364

Qy        121 LGM 123
              | |
Db        365 LDM 367


RESULT 3
Q63581
ID   Q63581      PRELIMINARY;      PRT;   430 AA.
AC   Q63581;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Rat T-kininogen (T-KG).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90034172; PubMed=2806908;
RA   Anderson K.P., Croyle M.L., Lingrel J.B.;
RT   "Primary structure of a gene encoding rat T-kininogen.";
RL   Gene 81:119-128(1989).
DR   EMBL; M29090; AAA42251.1; -.
DR   EMBL; M29083; AAA42251.1; JOINED.
DR   EMBL; M29084; AAA42251.1; JOINED.
DR   EMBL; M29091; AAA42251.1; JOINED.
DR   EMBL; M29085; AAA42251.1; JOINED.
DR   EMBL; M29086; AAA42251.1; JOINED.
DR   EMBL; M29087; AAA42251.1; JOINED.
DR   EMBL; M29088; AAA42251.1; JOINED.
DR   EMBL; M29089; AAA42251.1; JOINED.
DR   PIR; S68034; S68034.
DR   PIR; S68035; S68035.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 3.
DR   SMART; SM00043; CY; 3.
DR   PROSITE; PS00287; CYSTATIN; 2.
SQ   SEQUENCE   430 AA;  47618 MW;  45508DEF4BDC978C CRC64;

  Query Match             60.5%;  Score 399;  DB 11;  Length 430;
  Best Local Similarity   61.8%;  Pred. No. 6.3e-32;
  Matches   76;  Conservative   13;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |       | | ||||:|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFSLLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| ||||||:||| ||: ||| |: |||||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374


RESULT 4
Q9QWL5
ID   Q9QWL5      PRELIMINARY;      PRT;   167 AA.
AC   Q9QWL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Murine CMAP (CYSTATIN F) (LEUKOCYSTATIN).
GN   MURINE CMAP OR CST7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Morita M., Arakawa H., Yoshiuchi N.;
RT   "A novel cystatin-like metastasis associated gene.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
```

```
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AB015224; BAA34940.1; -.
DR   EMBL; AK004420; BAB23298.1; -.
DR   HSSP; P01034; 1G96.
DR   MGD; MGI:1298217; Cst7.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   167 AA;  18847 MW;  61F776D8445095FE CRC64;

  Query Match             26.5%;  Score 175;  DB 11;  Length 167;
  Best Local Similarity   34.4%;  Pred. No. 8.3e-10;
  Matches   42;  Conservative   23;  Mismatches   49;  Indels    8;  Gaps    4;

Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              ||  :   : :  | |: | ||:| : :   |:: | |   |  | ||  :| || ||||
Db         50 KDLI---SSVKPGFPKTIETNNPGVLKAARHSVEKFNNCTNDIFLFKESHVSKALVQVVK 106

Qy         62 GKKYFIDFVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVN 117
              | || ::     ||| |  : :| ::|:      | ::| | :||:|:||      | :
Db        107 GLKYMLEVKIGRTTCRKTMHHQL-DNCDFQTNPALKRTLYCYSEVWVIPWLHSFEVPVLL 165

Qy        118 CQ 119
              ||
Db        166 CQ 167


RESULT 5
Q7Z4J8
ID   Q7Z4J8      PRELIMINARY;      PRT;   167 AA.
AC   Q7Z4J8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin F (Leukocystatin).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT009825; AAP88827.1; -.
SQ   SEQUENCE   167 AA;  18857 MW;  E339025A5BD60177 CRC64;

  Query Match             25.1%;  Score 165.5;  DB 4;  Length 167;
  Best Local Similarity   32.5%;  Pred. No. 7.5e-09;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |   |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                ||| |   :  | : |:      | |:| | :||:|||| :    | : |
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 166


RESULT 6
```

```
Q7ZY91
ID   Q7ZY91      PRELIMINARY;      PRT;   462 AA.
AC   Q7ZY91;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to fetuin B.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC043891; AAH43891.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
SQ   SEQUENCE   462 AA;  53185 MW;  D7BAD339961739FB CRC64;

  Query Match             23.1%;  Score 152.5;  DB 13;  Length 462;
  Best Local Similarity   38.8%;  Pred. No. 5e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 63
              |  |   || | || : |:  |:| |    | : | ::| | |||||:::: | | | |
Db        142 PGVILSTCP-DCPTANEEITPTITETADTLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 200

Qy         64 KYFIDFVARETTCSKESNEELTESC 88
                |||  |  :|| | |     :  :|
Db        201 SYFIQFTIKETDCMKTQENVVLSNC 225


RESULT 7
Q7SYH2
ID   Q7SYH2      PRELIMINARY;      PRT;   462 AA.
AC   Q7SYH2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin domain fetuin-like protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ventral midgut;
RA   Costa R.M.B., Mason J., Lee M., Amaya E., Zorn A.M.;
RT   "Novel gene expression domains reveal early patterning of the Xenopus
RT   endoderm.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY260732; AAP82289.1; -.
SQ   SEQUENCE   462 AA;  53186 MW;  796F92774CC27721 CRC64;

  Query Match             23.1%;  Score 152.5;  DB 13;  Length 462;
  Best Local Similarity   38.8%;  Pred. No. 5e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 63
              |  |   || | || : |:  |:| |    | : | ::| | |||||:::: | | | |
Db        142 PGVILSTCP-DCPTANEEITPTITETADTLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 200

Qy         64 KYFIDFVARETTCSKESNEELTESC 88
                |||  |  :|| | |     :  :|
Db        201 SYFIQFTIKETDCMKTQENVVLSNC 225
```

```
RESULT 8
Q801E5
ID   Q801E5      PRELIMINARY;      PRT;   465 AA.
AC   Q801E5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical histidine-rich protein (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22480013; PubMed=12591597;
RA   Chen Y., Jurgens K., Hollemann T., Claussen M., Ramadori G.,
RA   Pieler T.;
RT   "Cell-autonomous and signal-dependent expression of liver and
RT   intestine marker genes in pluripotent precursor cells from Xenopus
RT   embryos.";
RL   Mech. Dev. 120:277-288(2003).
DR   EMBL; AY188284; AAO31610.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   465 AA;  53528 MW;  0B403AB4F78BBFD4 CRC64;

  Query Match             23.1%;  Score 152.5;  DB 13;  Length 465;
  Best Local Similarity   38.8%;  Pred. No. 5e-07;
  Matches   33;  Conservative   13;  Mismatches   34;  Indels    5;  Gaps    2;

Qy           8 PTKICVGCPRDIPTNSPELEETLTHT----ITKLNAENNATFYFKIDNVKKARVQVVAGK 63
               |  |    || | || : |:  |:| |     | : | ::| | |||||:::: | | |.|
Db         145 PGVILSTCP-DCPTANEEITPTITETAETLIAEYNKDSNNTRYFKIDHIERVRSQWVVGP 203

Qy          64 KYFIDFVARETTCSKESNEELTESC 88
                 ||| |  :|| | |      : ' :|
Db         204 SYFIQFTIKETDCMKTQENVVLSNC 228


RESULT 9
Q9EPX9
ID   Q9EPX9      PRELIMINARY;      PRT;   140 AA.
AC   Q9EPX9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cystatin C.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=21010502; PubMed=11144350;
RA   Taupin P.J., Ray J., Fischer W.H., Suhr S.T., Hakansson K., Grubb A.,
RA   Gage F.H.;
RT   "FGF-2-Responsive neural stem cell proliferation requires CCg, a novel
RT   autocrine/paracrine cofactor.";
RL   Neuron 28:385-397(2000).
DR   EMBL; AF311741; AAG40283.1; -.
DR   HSSP; P01034; 1G96.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
FT   CHAIN        21     140       CYSTATIN C.
FT   VARIANT      16      16       A -> G.
FT   VARIANT      84      84       L -> F.
SQ   SEQUENCE   140 AA;  15517 MW;  3A563406DD58D785 CRC64;

  Query Match             19.6%;  Score 129.5;  DB 11;  Length 140;
  Best Local Similarity   30.0%;  Pred. No. 2.6e-05;
  Matches   30;  Conservative   21;  Mismatches   44;  Indels    5;  Gaps    3;

Qy          13 VGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVAR 72
               :| | :   |   :   |   ::: |  :|  :: :   | :|| |:|||  ||:|
Db          30 LGAPEEADANEEGVRRALDFAVSEYNKGSNDAYHSRAIQVVRARKQLVAGVNYFLDVEMG 89

Qy          73 ETTCSKESNEELTESC---ETKKLGQSLDCNAEVYVVPWE 109
                 |||:| |   ||: |   :   | :   |: ::| |||:
Db          90 RTTCTK-SQTNLTD-CPFHDQPHLMRKALCSFQIYSVPWK 127


RESULT 10
Q800S8
ID   Q800S8      PRELIMINARY;      PRT;   455 AA.
AC   Q800S8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin-A.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Jia F.;
RT   "Danio rerio fetuin-A.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY217758; AAO61483.1; -.
DR   GO; GO:0005874; C:microtubule; IEA.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0007018; P:microtubule-based movement; IEA.
DR   InterPro; IPR002453; Beta_tubulin.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
SQ   SEQUENCE   455 AA;  50627 MW;  D822872926BA2ACB CRC64;

  Query Match             18.6%;  Score 122.5;  DB 13;  Length 455;
  Best Local Similarity   26.4%;  Pred. No. 0.00052;
  Matches   29;  Conservative   22;  Mismatches   44;  Indels   15;  Gaps    3;

Qy          10 KICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQ-VVAGKKYFID 68
               |  |   ||  :| : |:  |::   : | | ::|   |||:  | :   | :  |: ||
Db         140 KKCPDCPGLLPLHEPKALESVNAALAKFNKQSNHKSYFKLMEVGRISSQWMPMGQSYFTQ 199

Qy          69 FVARETTCSKESNEELTES-------------CETKKLG-QSLDCNAEVY 104
               |   || |:|:   :  |:             |:: |:| :  :   |:|
Db         200 FAIMETNCTKKDAPQNPEACKALCGDQATYGFCKSSKVGSEEPEVECEIY 249


RESULT 11
Q801Z5
ID   Q801Z5      PRELIMINARY;      PRT;   464 AA.
AC   Q801Z5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin long form.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
```

```
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tsai P.-L., Chang G.-D., Huang C.-J.;
RT   "Purification and cloning of carp fetuin.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY225965; AAO74862.1; -.
DR   GO; GO:0005874; C:microtubule; IEA.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0007018; P:microtubule-based movement; IEA.
DR   InterPro; IPR002453; Beta_tubulin.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00228; TUBULIN_B_AUTOREG; 1.
SQ   SEQUENCE   464 AA;  51698 MW;  7A54F71E44050895 CRC64;

  Query Match             17.9%;  Score 118;  DB 13;  Length 464;
  Best Local Similarity   24.2%;  Pred. No. 0.0015;
  Matches   31;  Conservative   27;  Mismatches   50;  Indels   20;  Gaps    5;

Qy         12 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQ-VVAGKKYFIDFV 70
              |  ||  :| : |:  |::   : | | |::   |||:  | :   | : :|: :|  |
Db        142 CPDCPGLLPLHDPKGLESVKTALQKFNKESDHKSYFKLMEVGRISTQWMFSGQSFFSQFA 201

Qy         71 ARETTCSKE---SNEELTES----------CETKKLG-QSLDCNAEVYVV-----PWEKK 111
                 || |: :    |||  ::          |:: |:| :  :   |:|        ||:
Db        202 IMETNCTNKEAPQNEESCKALCGEKARYGFCKSTKVGIEEPEVECEIYEAKNITHPWKHP 261

Qy        112 IYPTVNCQ 119
                :|:
Db        262 AQSRRDCK 269


RESULT 12
Q8CB17
ID   Q8CB17      PRELIMINARY;      PRT;   388 AA.
AC   Q8CB17;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fetuin beta.
GN   FETUB.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Vagina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK037043; BAC29682.1; -.
DR   MGD; MGI:1890221; Fetub.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR001363; Fetuin.
DR   Pfam; PF00031; cystatin; 2.
DR   SMART; SM00043; CY; 2.
DR   PROSITE; PS01254; FETUIN_1; 1.
DR   PROSITE; PS01255; FETUIN_2; 1.
SQ   SEQUENCE   388 AA;  42742 MW;  78CFAD73A8D8DC22 CRC64;

  Query Match             17.5%;  Score 115.5;  DB 11;  Length 388;
  Best Local Similarity   26.7%;  Pred. No. 0.0022;
  Matches   31;  Conservative   24;  Mismatches   48;  Indels   13;  Gaps    4;

Qy         12 CVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVA 71
              |  ||  |  ::|   |  | :: | |::: :  | ::  | ||  | |:|  |:::::
Db        154 CPDCPSPIDLSNPSALEAATESLAKFNSKSPSKKY-ELVKVTKAMNQWVSGPAYYVEYLI 212

Qy         72 RETTCSK--------ESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQ 119
                 :|  |:|          |: |     |:   :  ||        : | || :  || |:
Db        213 KEAPCTKSQASCSLQHSDSEPVGICQGSTVQSSL--RHVPLIQPVEKSV--TVTCE 264


RESULT 13
Q9NH95
ID   Q9NH95      PRELIMINARY;      PRT;   148 AA.
AC   Q9NH95;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ls-cystatin.
OS   Litomosoides sigmodontis.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC   Onchocercidae; Litomosoides.
OX   NCBI_TaxID=42156;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pfaff A.W., Hoffmann W.H., Taylor D.W., Schulz-Key H.;
RT   "Characterization and immunological properties of a cystein protease
RT   inhibitor of the filarial parasite Litomosoides sigmodontis.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF229173; AAF35896.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
DR   PROSITE; PS00287; CYSTATIN; 1.
FT   CHAIN        25    148       LS-CYSTATIN.
SQ   SEQUENCE   148 AA;  16686 MW;  2950AA89CA5339C9 CRC64;

  Query Match             17.2%;  Score 113.5;  DB 5;  Length 148;
  Best Local Similarity   34.1%;  Pred. No. 0.0011;
  Matches   30;  Conservative   16;  Mismatches   37;  Indels    5;  Gaps    3;

Qy         25 ELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESNEEL 84
              |::| |   :||:| ::|  ::     | |   |||| || ::     : | | |||::
Db         49 EIQEMLPSILTKVNQQSNDAYHLMPIKVLKVSSQVVAGMKYKMEIQVARSDCKKSSNEKI 108

Qy         85 -TESCETKKLGQSLD--CNAEVYVVPWE 109
               ::|  |||    |      ||:  ||
Db        109 DLKTC--KKLEGHPDQIITLEVWEKAWE 134


RESULT 14
Q80Y72
ID   Q80Y72      PRELIMINARY;      PRT;   140 AA.
AC   Q80Y72;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cystatin-like 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testicle;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testicle;
RA   Strausberg R.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC048646; AAH48646.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   InterPro; IPR003243; Cystatin_C/M.
DR   Pfam; PF00031; cystatin; 1.
DR   ProDom; PD001231; Cystatin_C/M; 1.
DR   SMART; SM00043; CY; 1.
SQ   SEQUENCE   140 AA;  16199 MW;  32633E99C4697DA0 CRC64;

  Query Match             17.1%;  Score 113;  DB 11;  Length 140;
  Best Local Similarity   29.4%;  Pred. No. 0.0012;
  Matches   25;  Conservative   18;  Mismatches   38;  Indels    4;  Gaps    2;

Qy         26 LEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESNEELT 85
              :  ||    |    |  :| |: :::  : : ::|:  | :| :       | | |  ||
Db         44 INSTLHFFIRSYNNASNDTYLYQVQKLIQGQMQLTTGVEYLVTVKIGRTKCKK--NETKK 101

Qy         86 ESC--ETKKLGQSLDCNAEVYVVPW 108
                ||  :: || :|| | : :| |||
Db        102 ASCPLQSSKLKKSLICKSLIYSVPW 126


RESULT 15
Q8K397
ID   Q8K397      PRELIMINARY;      PRT;   146 AA.
AC   Q8K397;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RIKEN cDNA 1110017E11 gene (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC027680; AAH27680.1; -.
DR   GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR   InterPro; IPR000010; Cystatin.
DR   Pfam; PF00031; cystatin; 1.
DR   SMART; SM00043; CY; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   146 AA;  16380 MW;  9D77BB9A6063A5C4 CRC64;

  Query Match             16.8%;  Score 111;  DB 11;  Length 146;
  Best Local Similarity   30.5%;  Pred. No. 0.002;
  Matches   29;  Conservative   18;  Mismatches   42;  Indels    6;  Gaps    4;

Qy         20 PTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKE 79
              ||: | :::     :   |  ::: :||:   |  |: |:||| ||::      | | |
Db         40 PTD-PRVQKAAQAAVASYNMGSDSLYYFRDTKVIDAKYQLVAGIKYYLTLDIESTECRKT 98

Qy         80 --SNEEL-TESCETKKLGQ--SLDCNAEVYVVPWE 109
                | | :   :|     ||   | || |:  |||:

Db         99 RVSGEHMDLTTCPLAAGGQQEKLRCNFELLEVPWK 133


Search completed: September 24, 2004, 14:10:17
Job time : 38.424 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:06:08 ; Search time 13.284 Seconds
(without alignments)
890.662 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	427	1	KGHUL1	kininogen, LMW pre
2	660	100.0	644	1	KGHUH1	kininogen, HMW pre
3	477	72.3	436	1	KGBOL1	kininogen, LMW I p
4	477	72.3	621	1	KGBOH1	kininogen, HMW I p
5	450	68.2	434	1	KGBOL2	kininogen, LMW II
6	450	68.2	619	1	KGBOH2	kininogen, HMW II
7	426	64.5	433	2	A28055	K-kininogen, LMW I
8	426	64.5	639	2	A25486	kininogen, HMW I p
9	409	62.0	430	2	A23897	major acute phase
10	409	62.0	430	2	B28055	T-kininogen, LMW I
11	402	60.9	423	1	KGRTM	major acute phase
12	401	60.8	430	1	KGRTT1	T-kininogen I prec
13	136	20.6	112	1	UDBO	cystatin - bovine
14	133	20.2	91	2	S68034	T-kininogen (clone
15	133	20.2	91	2	S68035	T-kininogen (clone
16	130	19.7	127	2	S07085	cystatin C precurs
17	129	19.5	120	2	S10587	cystatin C - rat
18	128	19.4	111	2	A28793	cystatin - puff ad
19	127.5	19.3	140	2	A36163	cystatin C precurs
20	127.5	19.3	146	1	UDHU	cystatin C precurs
21	125	18.9	141	2	B29632	cystatin SA precur
22	118.5	18.0	139	1	UDCH	cystatin precursor
23	113.5	17.2	122	2	A43644	sarcocystatin A pr
24	113	17.1	141	2	JQ1470	cystatin S precurs
25	111.5	16.9	111	1	JC2040	cystatin - chum sa
26	109.5	16.6	141	1	UDHUP2	cystatin SN precur
27	109	16.5	141	1	UDHUP1	cystatin S precurs
28	107.5	16.3	132	2	JC4918	cystatin precursor
29	107	16.2	139	2	T33740	hypothetical prote
30	105.5	16.0	162	2	A43428	onchocystatin - ne
31	97.5	14.8	135	2	JC4007	cystatin II - maiz
32	97	14.7	133	2	JC4536	cystatin precursor
33	95.5	14.5	143	2	T33301	hypothetical prote
34	95	14.4	142	2	A47142	cystatin D precurs
35	93.5	14.2	139	2	A45361	cystatin-related e
36	93.5	14.2	345	2	S21094	alpha-2-HS-glycopr
37	92	13.9	100	2	JH0269	cystatin - avocado
38	91.5	13.9	135	2	S27239	cysteine proteinas
39	90.5	13.7	375	2	A32827	fetuin precursor -
40	88.5	13.4	125	2	T00752	cysteine proteinas
41	88.5	13.4	134	2	S54828	cysteine proteinas
42	88.5	13.4	134	2	JC4882	cystatin - maize
43	85.5	13.0	367	1	WOHU	alpha-2-HS-glycopr
44	85	12.9	199	2	S65071	cystatin - field m
45	84.5	12.8	592	2	T50516	Ca2+ dependent lip

ALIGNMENTS

RESULT 1
KGHUL1
kininogen, LMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Homo sapiens (man)
C;Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C;Accession: A01280; B25276; A27900; A27699; A31905; A34030
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621; PMID:6441591
A;Accession: A01280
A;Molecule type: mRNA
A;Residues: 1-427 <OHK>
A;Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
A;Reference number: A92544; MUID:85234582; PMID:2989293
A;Accession: B25276
A;Molecule type: mRNA
A;Residues: 1-427 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, New
A;Title: Amino acid sequence of the light chain of human low molecular mass kininogen.
A;Reference number: A27900
A;Accession: A27900
A;Molecule type: protein
A;Residues: 390-427 <LOT>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021; PMID:3365237
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic flui
A;Reference number: A31905; MUID:89034061; PMID:3182782
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plas
A;Reference number: A34030; MUID:88106632; PMID:3337729
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>

R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583; PMID:2989294
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622; PMID:4952632
A;Contents: annotation; bradykinin
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (s
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
F;19-389,390-427/Product: LMW kininogen II #status predicted <MAT2>
F;19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-427/Product: LMW kininogen light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 660; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 3.3e-55;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375
```

RESULT 2
KGHUH1
kininogen, HMW precursor [validated] - human
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen; prokininogen
N;Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
C;Accession: A01279; A25276; S32422; A91153; A24871; A27899; A27699; A31905; A34030; S02
R;Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A;Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its identi
A;Reference number: A90490; MUID:85122621; PMID:6441591
A;Accession: A01279
A;Molecule type: mRNA
A;Residues: 1-389 <OHK>
A;Cross-references: GB:K02566; NID:g177889
R;Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A;Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low m
A;Reference number: A92544; MUID:85234582; PMID:2989293
A;Accession: A25276
A;Molecule type: mRNA
A;Residues: 1-592,'I',594-644 <TAK>
A;Cross-references: GB:M11437; NID:g186751; PIDN:AAB59550.1; PID:g386852
R;Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
FEBS Lett. 321, 93-97, 1993
A;Title: Cloning, expression and characterization of human kininogen domain 3.
A;Reference number: S32422; MUID:93223854; PMID:8467916
A;Accession: S32422
A;Molecule type: mRNA
A;Residues: 'ANSM',253-377 <AUE>
A;Note: differences are due to known cloning artifacts
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininoge
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Accession: A91153
A;Molecule type: protein
A;Residues: 379-644 <LOT>
A;Note: the bradykinin sequence preceding the light chain sequence was not determined in
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
Eur. J. Biochem. 154, 471-478, 1986
A;Title: Completion of the primary structure of human high-molecular-mass kininogen. The
A;Reference number: A24871; MUID:86108361; PMID:3484703
A;Accession: A24871
A;Molecule type: protein
A;Residues: 'Z',20-380 <KEL1>
R;Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New York
A;Title: Amino acid sequence of the light chain of human high molecular mass kininogen.
A;Reference number: A27899
A;Accession: A27899
A;Molecule type: protein
A;Residues: 379-389,'K',390-407,'Q',409-644 <KEL2>
R;Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A;Title: A new kinin moiety in human plasma kininogens.
A;Reference number: A27699; MUID:88209021; PMID:3365237
A;Accession: A27699
A;Molecule type: protein
A;Residues: 380-389 <MIN>
R;Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A;Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic fluid
A;Reference number: A31905; MUID:89034061; PMID:3182782
A;Accession: A31905
A;Molecule type: protein
A;Residues: 381-389 <MAE>
R;Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A;Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human plasm
A;Reference number: A34030; MUID:88106632; PMID:3337729
A;Accession: A34030
A;Molecule type: protein
A;Residues: 380-389 <SAS>
R;Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
A;Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and
A;Reference number: S02482; MUID:89076517; PMID:3264507
A;Accession: S02482
A;Molecule type: protein
A;Residues: 1-19;189-192;310-314;381-389 <LEN1>
R;Kato, H.; Matsumura, Y.; Maeda, H.
FEBS Lett. 232, 252-254, 1988
A;Title: Isolation and identification of hydroxyproline analogues of bradykinin in human
A;Reference number: A61495; MUID:88211869; PMID:3366244
A;Accession: A61495
A;Molecule type: protein
A;Residues: 380-389 <KAT1>
A;Experimental source: urine

A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: B61495
A;Molecule type: protein
A;Residues: 381-389 <KAT2>
A;Experimental source: urine
A;Note: this peptide had Pro-383 modified to 4-hydroxyproline
A;Accession: C61495
A;Molecule type: protein
A;Residues: 380-389 <KAT3>
R;Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A;Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A;Reference number: S14303; MUID:91192133; PMID:2013314
A;Accession: S14447
A;Molecule type: protein
A;Residues: 264-359,'N',361-375 <LEN2>
R;Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A;Title: Human mast cell tryptase isoforms: separation and examination of substrate-spec
A;Reference number: S55239; MUID:95251593; PMID:7733867
A;Accession: S55239
A;Molecule type: protein
A;Residues: 450-452,'X',454,'X',456 <LIT>
R;Straczek, J.; Maachi, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Bellevil
FEBS Lett. 373, 207-211, 1995
A;Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like
A;Reference number: S68059; MUID:96033974; PMID:7589467
A;Accession: S68059
A;Molecule type: protein
A;Residues: 431-434 <STR>
R;Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A;Title: Structural organization of the human kininogen gene and a model for its evoluti
A;Reference number: A92545; MUID:85234583; PMID:2989294
A;Contents: annotation; gene organization
R;Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A;Title: Structural features of plasma kinins and kininogens.
A;Reference number: A91455; MUID:90255622; PMID:4952632
A;Contents: annotation; bradykinin
C;Comment: The HMW kininogen precursor and the LMW form are produced from the same gene
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Genetics:
A;Gene: GDB:KNG
A;Cross-references: GDB:125256; OMIM:228960
A;Map position: 3q27-3q27
A;Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F;19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F;19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>
F;380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F;390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F;421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F;431-434/Product: low molecular weight growth promoting factor #status experimental <GF
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F;48/Binding site: carbohydrate (Asn) (covalent) #status absent
F;169,205,294/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F;383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;401,533,542,546,557,571,593,628/Binding site: carbohydrate (Thr) (covalent) #status ex
F;577/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 660; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 5.1e-55;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375
```

RESULT 3
KGBOL1
kininogen, LMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01283
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors an
A;Reference number: A93984; MUID:83117859; PMID:6572010
A;Accession: A01283
A;Molecule type: mRNA
A;Residues: 1-436 <NAW>
A;Cross-references: GB:J00010; GB:V00426; NID:g163256; PIDN:AAA30604.1; PID:g163257
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-436/Product: LMW kininogen I #status predicted <MAT>
F;19-378/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status predicted <BDY>
F;389-436/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27-406,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;47,87,168,169,197,204/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;378-379/Cleavage site: Met-Lys (kallikrein) #status predicted
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 72.3%; Score 477; DB 1; Length 436;
Best Local Similarity 71.9%; Pred. No. 9.2e-38;
Matches 87; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

```
Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              ||||||||::| |||: || :||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         62 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPL 121
              | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:|||||||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQPL 372

Qy        122 G 122
              |
Db        373 G 373
```

RESULT 4
KGBOH1
kininogen, HMW I precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01281; A91923; A91938; A29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininog
A;Reference number: A93317; MUID:84014106; PMID:6571699
A;Accession: A01281
A;Molecule type: mRNA
A;Residues: 1-621 <KIT>
A;Cross-references: GB:V01491; GB:K01757; NID:g491; PIDN:CAA24735.1; PID:g492
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420; PMID:4986212
A;Accession: A91923
A;Molecule type: protein
A;Residues: 378-393 <KAT>
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265; PMID:1169237
A;Accession: A91938
A;Molecule type: protein
A;Residues: 458-498 <HAN>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530; PMID:3546295
A;Accession: A29559
A;Molecule type: protein
A;Residues: 'Z',20-123,'I',125-127,'I',129-378 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-621/Product: HMW prokininogen I #status predicted <MAT>
F;19-379/Product: HMW kininogen I heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;379-388/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;380-388/Product: bradykinin (kallidin I) #status experimental <BDY>
F;389-621/Product: HMW kininogen I light chain #status experimental <LCH>
F;417-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-591,82-93,106-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds:
F;87,168,169,204/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;378-379/Cleavage site: Met-Lys (kallikrein) #status experimental
F;382/Modified site: 4-hydroxyproline (Pro) #status predicted
F;388-389/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;398,406,512/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;399,400,520,524,536,548,553,570/Binding site: carbohydrate (Thr) (covalent) #status exp
F;498-499/Cleavage site: Arg-Thr (kallikrein) #status experimental

Query Match 72.3%; Score 477; DB 1; Length 621;
Best Local Similarity 71.9%; Pred. No. 1.4e-37;
Matches 87; Conservative 14; Mismatches 20; Indels 0; Gaps 0;

```
Qy          2 KDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVA 61
              |||||||::| |||: || :.||:||| |:|:| |||||::  |||||| |||| |||||
Db        253 KDFVQPPTRLCAGCPKPIPVDSPDLEEPLSHSIAKLNAEHDGAFYFKIDTVKKATVQVVA 312

Qy         62 GKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPL 121
              | || | |:|||||||| ||||||:|||     || | |:| |||||||:|:|||||||||
Db        313 GLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQPL 372

Qy        122 G 122
              |
Db        373 G 373
```

RESULT 5
KGBOL2
kininogen, LMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C;Accession: A01284
R;Nawa, H.; Kitamura, N.; Hirose, T.; Asai, M.; Inayama, S.; Nakanishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 90-94, 1983
A;Title: Primary structures of bovine liver low molecular weight kininogen precursors and
A;Reference number: A93984; MUID:83117859; PMID:6572010
A;Accession: A01284
A;Molecule type: mRNA
A;Residues: 1-434 <NAW>
A;Cross-references: GB:V00427; GB:J00011; NID:g489; PIDN:CAA23710.1; PID:g490
C;Comment: The LMW kininogen precursor is produced from the same gene as the HMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, ir
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyco
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-434/Product: LMW kininogen II #status predicted <MAT>
F;19-377/Product: LMW kininogen I heavy chain #status predicted <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status predicted <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status predicted <BDY>
F;387-434/Product: LMW kininogen I light chain #status experimental <LCH>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;27-404,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bonds: #
F;47,87,168,169,197,204,280/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;376-377/Cleavage site: Met-Lys (kallikrein) #status predicted
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status predicted

Query Match 68.2%; Score 450; DB 1; Length 434;
Best Local Similarity 68.9%; Pred. No. 3.4e-35;
Matches 84; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |:||:  |  :|||||: || :||:||| | |:| |||||:: |||||||  |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              | || | |:|||||||| ||||||:|||     || | |:| |||||||:|:|||||||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQP 369

Qy        121 LG 122
              ||
```

```
Db        370 LG 371


RESULT 6
KGBOH2
kininogen, HMW II precursor - bovine
N;Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N;Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
C;Accession: A01282; A91923; A91941; A91938; B29559
R;Kitamura, N.; Takagaki, Y.; Furuto, S.; Tanaka, T.; Nawa, H.; Nakanishi, S.
Nature 305, 545-549, 1983
A;Title: A single gene for bovine high molecular weight and low molecular weight kininog
A;Reference number: A93317; MUID:84014106; PMID:6571699
A;Accession: A01282
A;Molecule type: mRNA
A;Residues: 1-619 <KIT>
A;Cross-references: GB:V01492; GB:K01758; NID:g493; PIDN:CAA24736.1; PID:g494
R;Kato, H.; Nagasawa, S.; Suzuki, T.
J. Biochem. 67, 313-323, 1970
A;Title: Studies on the structure of bovine kininogen: cleavages of disulfide bonds and
A;Reference number: A91923; MUID:70180420; PMID:4986212
A;Accession: A91923
A;Molecule type: protein
A;Residues: 376-391 <KAT>
R;Han, Y.N.; Kato, H.; Iwanaga, S.; Suzuki, T.
J. Biochem. 79, 1201-1222, 1976
A;Title: Primary structure of bovine plasma high-molecular-weight kininogen. The amino a
A;Reference number: A91941; MUID:76260155; PMID:956151
A;Accession: A91941
A;Molecule type: protein
A;Residues: 387-455 <HAN>
A;Note: 398-Pro, 401-Val, and 455-Lys were also found
R;Han, Y.N.; Komiya, M.; Iwanaga, S.; Suzuki, T.
J. Biochem. 77, 55-68, 1975
A;Title: Studies on the primary structure of bovine high-molecular-weight kininogen. Ami
A;Reference number: A91938; MUID:75170265; PMID:1169237
A;Accession: A91938
A;Molecule type: protein
A;Residues: 456-496 <HA2>
R;Sueyoshi, T.; Miyata, T.; Hashimoto, N.; Kato, H.; Hayashida, H.; Miyata, T.; Iwanaga,
J. Biol. Chem. 262, 2768-2779, 1987
A;Title: Bovine high molecular weight kininogen. The amino acid sequence, positions of c
A;Reference number: A92627; MUID:87137530; PMID:3546295
A;Accession: B29559
A;Molecule type: protein
A;Residues: 'Z',20-104,'E',106-256,'XX',257-376 <SUE>
R;Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
Eur. J. Biochem. 152, 307-314, 1985
A;Title: The amino acid sequence of the light chain of human high-molecular-mass kininog
A;Reference number: A91153; MUID:86030270; PMID:4054110
A;Contents: annotation; bovine cleavage sites; bovine carbohydrate binding sites
R;Sueyoshi, T.; Miyata, T.; Kato, H.; Iwanaga, S.
Seikagaku 56, 808, 1984
A;Title: Disulfide bonds in bovine HMW kininogens.
A;Reference number: A94300
A;Contents: annotation; disulfide bonds
A;Note: article in Japanese
C;Comment: The HMW kininogen precursor is produced from the same gene as the LMW form as
C;Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
C;Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is impo
C;Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, i
xyproline residue is present in the kininogen prior to the release of bradykinin.
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; dupl
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-619/Product: HMW kininogen II #status predicted <MAT>
F;19-376/Product: HMW kininogen II heavy chain #status experimental <HCH>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;261-372/Domain: cystatin homology <CY3>
F;377-386/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F;378-386/Product: bradykinin (kallidin I) #status experimental <BDY>
F;387-619/Product: HMW kininogen II light chain #status experimental <LCH>
F;418-488/Region: glycine/histidine/lysine-rich
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27-589,82-93,106-125,141-144,205-217,228-247,261-264,325-337,348-367/Disulfide bonds:
F;47/Binding site: carbohydrate (Asn) (covalent) #status absent
F;87,168,169,204,280/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;136/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F;197/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;376-377/Cleavage site: Met-Lys (kallikrein) #status experimental
F;380/Modified site: 4-hydroxyproline (Pro) #status predicted
F;386-387/Cleavage site: Arg-Ser (kallikrein) #status experimental
F;396,400,404,510/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;397,398,518,522,534,546,551,568/Binding site: carbohydrate (Thr) (covalent) #status ex
F;496-497/Cleavage site: Arg-Thr (kallikrein) #status experimental

  Query Match             68.2%;  Score 450;  DB 1;  Length 619;
  Best Local Similarity   68.9%;  Pred. No. 5.1e-35;
  Matches   84;  Conservative   14;  Mismatches   22;  Indels    2;  Gaps    1;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              |:||:  |  :|||||: || :||:||| | |:| |||||:: ||||||| |||| ||||
Db        252 GEDFL--PPMVCVGCPKPIPVDSPDLEEALNHSIAKLNAEHDGTFYFKIDTVKKATVQVV 309

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
               | || | |:|||||||| ||||||:|||    || | |:| |||||||:|:||||||||
Db        310 GGLKYSIVFIARETTCSKGSNEELTKSCEINIHGQILHCDANVYVVPWEEKVYPTVNCQP 369

Qy        121 LG 122
              ||
Db        370 LG 371


RESULT 7
A28055
K-kininogen, LMW I precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C;Accession: A28055
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: A28055
A;Molecule type: mRNA
A;Residues: 1-433 <FUR>
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-433/Product: K-kininogen, LMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

  Query Match             64.5%;  Score 426;  DB 2;  Length 433;
  Best Local Similarity   65.0%;  Pred. No. 6.6e-33;
  Matches   80;  Conservative   14;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   : | ||||:|| :||||:| | |:| :|||||| ||||||| |||| |||
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||||| ||| |||||:|||||:||| ||: ||| |: ||| |:
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTVKCKV 372

Qy        121 LGM 123
              |||
Db        373 LDM 375


RESULT 8
```

A25486
kininogen, HMW I precursor - rat
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C;Accession: A25486
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: A25486
A;Molecule type: mRNA
A;Residues: 1-639 <KIT>
A;Note: the authors translated the codon CAA for residue 347 as Asn
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-639/Product: kininogen, HMW I #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;142-253/Domain: cystatin homology <CY2>
F;264-375/Domain: cystatin homology <CY3>

Query Match 64.5%; Score 426; DB 2; Length 639;
Best Local Similarity 65.0%; Pred. No. 1e-32;
Matches 80; Conservative 14; Mismatches 29; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   : | ||||:|| :||||:| | |:| :|||||| ||||||| |||| |||
Db        253 GDDLFELLPEDCPGCPRNIPVDSPELKEALGHSIAQLNAENNHTFYFKIDTVKKATSQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||||| ||| |||||:|||||:||| ||: ||| |: ||| |:
Db        313 AGTKYVIEFIARETKCSKESNAELTADCETKRLGQSLNCNANVYMRPWENKVVPTVKCKV 372

Qy        121 LGM 123
              | |
Db        373 LDM 375
```

RESULT 9
A23897
major acute phase alpha-1 protein (version 2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 20-Aug-1999
C;Accession: A23897; B23897
R;Anderson, K.P.; Heath, E.C.
J. Biol. Chem. 260, 12065-12071, 1985
A;Title: The relationship between rat major acute phase protein and the kininogens.
A;Reference number: A23897; MUID:86008266; PMID:2413019
A;Accession: A23897
A;Molecule type: protein
A;Residues: 1-14 <AND1>
A;Accession: B23897
A;Molecule type: mRNA
A;Residues: 5-430 <AND2>
A;Cross-references: GB:M11661; NID:g205307; PIDN:AAA41570.1; PID:g205308
A;Note: the authors translated the codon CTC for residue 410 as Arg, CTA for residue 415
C;Superfamily: kininogen; cystatin homology
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>

Query Match 62.0%; Score 409; DB 2; Length 430;
Best Local Similarity 61.8%; Pred. No. 2.7e-31;
Matches 76; Conservative 15; Mismatches 32; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |      | | |||::|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||::| |||  |||| |||||:||| ||: ||| |: ||| ||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374
```

RESULT 10
B28055
T-kininogen, LMW II precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 12-Dec-1997
C;Accession: B28055; E25486; B28526; C28526
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and T
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: B28055
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: E25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of ac
A;Reference number: A92729; MUID:88087226; PMID:3121623
A;Accession: B28526
A;Molecule type: protein
A;Residues: 'E',20-25,'MD',28-48,376-430 <ENJ>
A;Accession: C28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-388,'R',390-419,'ER',422-430 <EN2>
C;Superfamily: kininogen; cystatin homology
C;Keywords: glycoprotein; pyroglutamic acid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen, LMW II #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experiment
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #status

Query Match 62.0%; Score 409; DB 2; Length 430;
Best Local Similarity 61.8%; Pred. No. 2.7e-31;
Matches 76; Conservative 15; Mismatches 32; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |      | | |||::|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFSLLPKKCFGCPKNIPVDSPELKEALGHSIAQLNAQHNHLFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              || || |:|:|||| |||::| |||  |||| |||||:||| ||: ||| |: ||| ||
Db        312 AGTKYVIEFIARETNCSKQTNTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374
```

RESULT 11
KGRTM
major acute phase alpha-1 protein precursor - rat (fragment)
N;Contains: bradykinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 12-Apr-1996
C;Accession: A01285
R;Cole, T.; Inglis, A.S.; Roxburgh, C.M.; Howlett, G.J.; Schreiber, G.

FEBS Lett. 182, 57-61, 1985
A;Title: Major acute phase alpha1-protein of the rat is homologous to bovine kininogen a
A;Reference number: A01285; MUID:85127561; PMID:2578992
A;Accession: A01285
A;Molecule type: mRNA
A;Residues: 1-423 <COL>
C;Comment: This plasma glycoprotein inhibits cysteine proteinases. During acute inflamma
r.
C;Superfamily: kininogen; cystatin homology
C;Keywords: bradykinin; cysteine proteinase inhibitor; duplication; glycoprotein; inflam
F;1-11/Domain: signal sequence (fragment) #status predicted <SIG>
F;12-423/Product: major acute phase alpha-1 protein #status predicted <MAT>
F;12-123/Domain: cystatin homology <CY1>
F;134-245/Domain: cystatin homology <CY2>
F;256-367/Domain: cystatin homology <CY3>
F;371-379/Product: bradykinin #status predicted <BDY>
F;12/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;161,197/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.9%; Score 402; DB 1; Length 423;
Best Local Similarity 61.8%; Pred. No. 1.2e-30;
Matches 76; Conservative 14; Mismatches 33; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||| |||| |||
Db        245 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 304

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| ||||||||| ||: ||| |: ||| ||
Db        305 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 364

Qy        121 LGM 123
              | |
Db        365 LDM 367
```

RESULT 12
KGRTT1
T-kininogen I precursor - rat
N;Alternate names: 73K protein; LMW kininogen T-I
N;Contains: bradykinin; T-kinin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 22-Jun-1999
C;Accession: A01286; D25486; A28526; PL0193; JQ0027; B25488; A28525; S68036
R;Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A;Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin and
nhibitor.
A;Reference number: A92496; MUID:86008264; PMID:2413018
A;Accession: A01286
A;Molecule type: mRNA
A;Residues: 1-430 <FUR>
A;Cross-references: GB:M11883; NID:g205084; PIDN:AAA41489.1; PID:g205085
R;Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A;Title: Differing expression patterns and evolution of the rat kininogen gene family.
A;Reference number: A92625; MUID:87137443; PMID:3029068
A;Accession: D25486
A;Molecule type: DNA
A;Residues: 375-430 <KIT>
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 973-979, 1988
A;Title: Purification and characterization of rat T-kininogens isolated from plasma of a
A;Reference number: A92729; MUID:88087226; PMID:3121623
A;Accession: A28526
A;Molecule type: protein
A;Residues: 'E',20-48;376-430 <ENJ>
R;Kanda, S.; Sugiyama, K.; Takahashi, M.; Shumiya, S.; Tomino, S.; Nagase, S.
Jpn. J. Cancer Res. 81, 63-68, 1990
A;Title: Identification of a protein increasing in serum of Nagase analbuminemic rats be
A;Reference number: PL0193; MUID:90216390; PMID:2108948
A;Accession: PL0193
A;Molecule type: mRNA
A;Residues: 330-420,'R',422-429,'P' <KAN>
R;Anderson, K.P.; Croyle, M.L.; Lingrel, J.B.
Gene 81, 119-128, 1989
A;Title: Primary structure of a gene encoding rat T-kininogen.
A;Reference number: JQ0027; MUID:90034172; PMID:2806908
A;Accession: JQ0027
A;Molecule type: DNA
A;Residues: 1-60,'E',62-113,'R',115-165,'F',167-178,'TKI',182-211,'F',213-256,'S',258-388
A;Experimental source: strain Sprague-Dawley
R;Kageyama, R.; Kitamura, N.; Ohkubo, H.; Nakanishi, S.
J. Biol. Chem. 262, 2345-2351, 1987
A;Title: Differing utilization of homologous transcription initiation sites of rat K and
A;Reference number: A25488; MUID:87137465; PMID:3818598
A;Accession: B25488
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <KAG>
A;Cross-references: GB:M14356; NID:g205090; PIDN:AAA41492.1; PID:g205091
R;Enjyoji, K.; Kato, H.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biol. Chem. 263, 965-972, 1988
A;Title: Purification and characterization of two kinds of low molecular weight kininoger
A;Reference number: A28525; MUID:88087225; PMID:3335530
A;Accession: A28525
A;Molecule type: protein
A;Residues: 376-430 <EN2>
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of agi
A;Reference number: S68034; MUID:96032652; PMID:7574705
A;Accession: S68036
A;Molecule type: mRNA
A;Residues: 340-430 <SIE>
A;Experimental source: clone pSG17
C;Comment: At least three types of LMW kininogen precursors are present in rat plasma, t
ceding bradykinin.
C;Comment: T-kininogens contain T-kinin (I-S-bradykinin), a novel kinin isolated after t
d of an Arg or Lys, it is probably not released from its precursor by either tissue or p
C;Comment: The T-kininogens are produced in response to an inflammatory stimulant.
C;Genetics:
A;Introns: 65/3; 102/3; 130/1; 187/3; 223/2; 252/1; 309/3; 345/3; 374/3; 398/3
C;Superfamily: kininogen; cystatin homology
C;Keywords: acute phase; bradykinin; cysteine proteinase inhibitor; duplication; glycopr
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-430/Product: T-kininogen I #status experimental <MAT>
F;19-130/Domain: cystatin homology <CY1>
F;141-252/Domain: cystatin homology <CY2>
F;263-374/Domain: cystatin homology <CY3>
F;378-386/Product: bradykinin #status predicted <BDY>
F;19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;82,126,168,204,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;83-94,107-125,141-144,205-217,228-247,263-266,327-339,350-369/Disulfide bonds: #status

Query Match 60.8%; Score 401; DB 1; Length 430;
Best Local Similarity 61.8%; Pred. No. 1.6e-30;
Matches 76; Conservative 14; Mismatches 33; Indels 0; Gaps 0;

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||| |||| |||
Db        252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| ||||||:||| ||: ||| |: ||| ||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374
```

RESULT 13
UDBO
cystatin - bovine
N;Alternate names: thiol proteinase inhibitor

```
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 06-Dec-1996
C;Accession: A01271
R;Hirado, M.; Tsunasawa, S.; Sakiyama, F.; Niinobe, M.; Fujii, S.
FEBS Lett. 186, 41-45, 1985
A;Title: Complete amino acid sequence of bovine colostrum low-M-r cysteine proteinase in
A;Reference number: A01271; MUID:85231205; PMID:3891407
A;Accession: A01271
A;Molecule type: protein
A;Residues: 1-112 <HIR>
C;Superfamily: cystatin; cystatin homology
C;Keywords: colostrum; cysteine proteinase inhibitor
F;2-112/Domain: cystatin homology <CYS>
F;48-52/Region: inhibitory #status predicted
F;66-76,90-110/Disulfide bonds: #status predicted

  Query Match             20.6%;  Score 136;  DB 1;  Length 112;
  Best Local Similarity   30.6%;  Pred. No. 6.4e-06;
  Matches   34;  Conservative   22;  Mismatches   31;  Indels   24;  Gaps    5;

Qy         22 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 81
              |   ::| |:  ::: |  :|  :  ::  | :|| |||:|  ||:|     |||:|  :
Db         12 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 69

Qy         82 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVN-----CQ 119
              :   :||        : :||     |: :||||||      |:|     ||
Db         70 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 111


RESULT 14
S68034
T-kininogen (clone pSG22) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 07-May-1999
C;Accession: S68034
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of agi
A;Reference number: S68034; MUID:96032652; PMID:7574705
A;Accession: S68034
A;Molecule type: mRNA
A;Residues: 1-91 <SIE>
A;Experimental source: liver
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; cysteine proteinase inhibitor; plasma

  Query Match             20.2%;  Score 133;  DB 2;  Length 91;
  Best Local Similarity   68.6%;  Pred. No. 9.8e-06;
  Matches   24;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         89 ETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 123
              ||| |||||:||| ||: ||| |: ||| || | |
Db          1 ETKHLGQSLNCNANVYMRPWENKVVPTVRCQALDM 35


RESULT 15
S68035
T-kininogen (clone pSG17) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 07-May-1999
C;Accession: S68035
R;Sierra, F.; Walter, R.; Vautravers, P.; Guigoz, Y.
Arch. Biochem. Biophys. 322, 333-338, 1995
A;Title: Identification of several isoforms of T-kininogen expressed in the liver of agi
A;Reference number: S68034; MUID:96032652; PMID:7574705
A;Accession: S68035
A;Molecule type: mRNA
A;Residues: 1-91 <SIE>
A;Experimental source: liver
C;Superfamily: kininogen; cystatin homology
C;Keywords: alternative splicing; cysteine proteinase inhibitor; plasma

  Query Match             20.2%;  Score 133;  DB 2;  Length 91;
  Best Local Similarity   68.6%;  Pred. No. 9.8e-06;
  Matches   24;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         89 ETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQPLGM 123
              ||| |||||:||| ||: ||| |: ||| || | |
Db          1 ETKHLGQSLNCNANVYMRPWENKVVPTVRCQALDM 35


Search completed: September 24, 2004, 14:10:48
Job time : 14.284 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 14:08:41 ; Search time 43.296 Seconds
(without alignments)
913.519 Million cell updates/sec

Title: US-10-661-784-1
Perfect score: 660
Sequence: 1 GKDFVQPPTKICVGCPRDIP.....YVVPWEKKIYPTVNCQPLGM 123

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	660	100.0	390	15	US-10-162-335-82	Sequence 82, Appl
2	660	100.0	398	15	US-10-162-335-70	Sequence 70, Appl
3	660	100.0	427	10	US-09-919-039-29	Sequence 29, Appl
4	660	100.0	615	15	US-10-162-335-72	Sequence 72, Appl
5	660	100.0	644	15	US-10-162-335-74	Sequence 74, Appl
6	660	100.0	644	15	US-10-162-335-84	Sequence 84, Appl
7	402	60.9	424	14	US-10-316-253-217	Sequence 217, App
8	401	60.8	430	14	US-10-316-253-215	Sequence 215, App
9	169	25.6	178	9	US-09-969-834-1	Sequence 1, Appli
10	165.5	25.1	145	14	US-10-329-428-2	Sequence 2, Appli
11	165.5	25.1	167	10	US-09-746-783-197	Sequence 197, App
12	138.5	21.0	121	9	US-09-775-932-14	Sequence 14, Appl
13	138.5	21.0	128	9	US-09-775-932-12	Sequence 12, Appl
14	138.5	21.0	149	9	US-09-940-497-2	Sequence 2, Appli
15	136	20.6	112	8	US-08-849-303-16	Sequence 16, Appl
16	136	20.6	112	16	US-10-655-136-16	Sequence 16, Appl
17	135	20.5	118	9	US-09-775-932-24	Sequence 24, Appl
18	130	19.7	127	8	US-08-849-303-19	Sequence 19, Appl
19	130	19.7	127	16	US-10-655-136-19	Sequence 19, Appl
20	129.5	19.6	140	14	US-10-376-564-46	Sequence 46, Appl
21	129.5	19.6	140	14	US-10-376-564-48	Sequence 48, Appl
22	129.5	19.6	317	12	US-10-210-172-82	Sequence 82, Appl
23	129.5	19.6	345	12	US-10-210-172-86	Sequence 86, Appl
24	129.5	19.6	356	12	US-10-210-172-84	Sequence 84, Appl
25	129.5	19.6	369	12	US-10-210-172-78	Sequence 78, Appl
26	129.5	19.6	369	12	US-10-210-172-80	Sequence 80, Appl
27	129.5	19.6	382	12	US-10-315-664-93	Sequence 93, Appl
28	129.5	19.6	382	12	US-09-978-360A-425	Sequence 425, App
29	128	19.4	111	8	US-08-849-303-26	Sequence 26, Appl
30	128	19.4	111	16	US-10-655-136-26	Sequence 26, Appl
31	127.5	19.3	120	9	US-09-775-932-2	Sequence 2, Appli
32	127.5	19.3	120	16	US-10-695-194-2	Sequence 2, Appli
33	127.5	19.3	140	8	US-08-849-303-18	Sequence 18, Appl
34	127.5	19.3	140	16	US-10-655-136-18	Sequence 18, Appl
35	127.5	19.3	146	8	US-08-849-303-17	Sequence 17, Appl
36	127.5	19.3	146	9	US-09-940-497-3	Sequence 3, Appli
37	127.5	19.3	146	9	US-09-969-834-3	Sequence 3, Appli
38	127.5	19.3	146	14	US-10-329-428-3	Sequence 3, Appli
39	127.5	19.3	146	14	US-10-376-564-47	Sequence 47, Appl
40	127.5	19.3	146	16	US-10-655-136-17	Sequence 17, Appl
41	127.5	19.3	146	16	US-10-695-194-1	Sequence 1, Appli
42	127.5	19.3	249	16	US-10-257-384A-4	Sequence 4, Appli
43	127.5	19.3	641	16	US-10-257-384A-2	Sequence 2, Appli
44	125	18.9	121	9	US-09-775-932-8	Sequence 8, Appli
45	125	18.9	141	8	US-08-849-303-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-10-162-335-82
; Sequence 82, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04

```
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 82
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-335-82

  Query Match             100.0%;  Score 660;  DB 15;  Length 390;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 275

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 335

Qy        121 LGM 123
              |||
Db        336 LGM 338


RESULT 2
US-10-162-335-70
; Sequence 70, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 70
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-335-70

  Query Match             100.0%;  Score 660;  DB 15;  Length 398;
  Best Local Similarity   100.0%;  Pred. No. 3e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 343

Qy        121 LGM 123
              |||
Db        344 LGM 346


RESULT 3
US-09-919-039-29
; Sequence 29, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
;  APPLICANT: Kaser, Matthew R.
;  TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
;  FILE REFERENCE: PA-0035 US
;  CURRENT APPLICATION NUMBER: US/09/919,039
;  CURRENT FILING DATE:  2002-09-09
;  PRIOR APPLICATION NUMBER: 60/222,113
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 401
;  SOFTWARE: PERL Program
; SEQ ID NO 29
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030108871A1 167507CD1
US-09-919-039-29

  Query Match             100.0%;  Score 660;  DB 10;  Length 427;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 4
US-10-162-335-72
; Sequence 72, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 72
;   LENGTH: 615
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-335-72

  Query Match           100.0%;  Score 660;  DB 15;  Length 615;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 283

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 343

Qy        121 LGM 123
              |||
Db        344 LGM 346


RESULT 5
US-10-162-335-74
; Sequence 74, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
```

```
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 74
;    LENGTH: 644
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-162-335-74

  Query Match             100.0%;  Score 660;  DB 15;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 6
US-10-162-335-84
; Sequence 84, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David W.
;  APPLICANT:  Baumgartner, Jason C.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Edinger, Shlomit R.
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Guo, Xiaojia (Sasha)
;  APPLICANT:  Hjalt, Tord
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Li, Li
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Pena, Carol E. A.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Vernet, Corine A. M.
;  APPLICANT:  Voss, Edward Z.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
;  FILE REFERENCE: 21402-377 B
;  CURRENT APPLICATION NUMBER: US/10/162,335
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,607
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/295,661
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/296,404
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/296,418
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: 60/297,414
;  PRIOR FILING DATE: 2001-06-11
;  PRIOR APPLICATION NUMBER: 60/297,567
;  PRIOR FILING DATE: 2001-06-12
;  PRIOR APPLICATION NUMBER: 60/298,285
;  PRIOR FILING DATE: 2001-06-14
;  PRIOR APPLICATION NUMBER: 60/298,556
;  PRIOR FILING DATE: 2001-06-15
;  PRIOR APPLICATION NUMBER: 60/299,949
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: 60/300,883
;  PRIOR FILING DATE: 2001-06-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 84
;    LENGTH: 644
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-162-335-84

  Query Match             100.0%;  Score 660;  DB 15;  Length 644;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-64;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 312

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 372

Qy        121 LGM 123
              |||
Db        373 LGM 375


RESULT 7
US-10-316-253-217
; Sequence 217, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Peters, Kevin
;  APPLICANT:  Thompson, Larry
;  APPLICANT:  Wang, Feng
;  APPLICANT:  Greis, Kenneth
;  TITLE OF INVENTION: Angiogenesis Modulating Proteins
;  FILE REFERENCE: 8865M
;  CURRENT APPLICATION NUMBER: US/10/316,253
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: US 60/355,295
;  PRIOR FILING DATE: 2002-02-08
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 217
;    LENGTH: 424
;    TYPE: PRT
;    ORGANISM: Rattus norvegicus
US-10-316-253-217

  Query Match             60.9%;  Score 402;  DB 14;  Length 424;
  Best Local Similarity   61.8%;  Pred. No. 1.1e-35;
  Matches   76;  Conservative   14;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  ||||||  ||||  |||
Db        246 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIARLNAQHNHIFYFKIDTVKKATSQVV 305

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: ||| ||
Db        306 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 365

Qy        121 LGM 123
              | |
Db        366 LDM 368
```

```
RESULT 8
US-10-316-253-215
; Sequence 215, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Peters, Kevin
;  APPLICANT:  Thompson, Larry
;  APPLICANT:  Wang, Feng
;  APPLICANT:  Greis, Kenneth
;  TITLE OF INVENTION: Angiogenesis Modulating Proteins
;  FILE REFERENCE: 8865M
;  CURRENT APPLICATION NUMBER: US/10/316,253
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: US 60/355,295
;  PRIOR FILING DATE: 2002-02-08
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 215
;   LENGTH: 430
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-10-316-253-215

  Query Match             60.8%;  Score 401;  DB 14;  Length 430;
  Best Local Similarity   61.8%;  Pred. No. 1.5e-35;
  Matches   76;  Conservative   14;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 GKDFVQPPTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVV 60
              | |  :   | | ||||:|| :||||:| | |:| :|||::|  |||||| ||||  |||
Db        252 GDDLFELLPKNCRGCPREIPVDSPELKEALGHSIAQLNAQHNHIFYFKIDTVKKATSQVV 311

Qy         61 AGKKYFIDFVARETTCSKESNEELTESCETKKLGQSLDCNAEVYVVPWEKKIYPTVNCQP 120
              ||  | |:|:|||| |||:|  |||  |||| |||||:||| ||: ||| |: ||| ||
Db        312 AGVIYVIEFIARETNCSKQSKTELTADCETKHLGQSLNCNANVYMRPWENKVVPTVRCQA 371

Qy        121 LGM 123
              | |
Db        372 LDM 374


RESULT 9
US-09-969-834-1
; Sequence 1, Application US/09969834
; Patent No. US20020102711A1
;   GENERAL INFORMATION:
;        APPLICANT: Bandman, Olga
;                   Goli, Surya K.
;        TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
;                            PROTEIN
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/969,834
;             FILING DATE: 01-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/471,765
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US/08/791,522
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/471,765
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0193 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-855-0555
;             TELEFAX: 415-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 178 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             CLONE: 30443
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-969-834-1

  Query Match             25.6%;  Score 169;  DB 9;  Length 178;
  Best Local Similarity   32.5%;  Pred. No. 2.1e-10;
  Matches   39;  Conservative   22;  Mismatches   49;  Indels   10;  Gaps    4;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKIYPTVN--CQPLGM 123
                 ||| |  :  | : |:      | |:| | :||:|||      |  :   | || :
Db        114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPW----VPALRGACSPLSL 168


RESULT 10
US-10-329-428-2
; Sequence 2, Application US/10329428
; Publication No. US20030114646A1
; GENERAL INFORMATION:
;  APPLICANT: Li, et al.
;  TITLE OF INVENTION: Human Cystatin F
;  FILE REFERENCE: PF265P1D2
;  CURRENT APPLICATION NUMBER: US/10/329,428
;  CURRENT FILING DATE:  2002-12-27
;  PRIOR APPLICATION NUMBER: 60/014,795
;  PRIOR FILING DATE: 1996-04-03
;  PRIOR APPLICATION NUMBER: 08/832,535
;  PRIOR FILING DATE: 1997-04-03
;  PRIOR APPLICATION NUMBER: 09/019,485
;  PRIOR FILING DATE: 1998-01-29
;  PRIOR APPLICATION NUMBER: 09/528,436
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 145
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-329-428-2

  Query Match             25.1%;  Score 165.5;  DB 14;  Length 145;
  Best Local Similarity   32.5%;  Pred. No. 3.9e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy          9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
              :::  | |: | || | : :   ::: | |    |  | ||   : :| ||:| | || ::
Db         32 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 91

Qy         69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                 ||| |  :  | : |:      | |:| | :||:|||| :    | : |
Db         92 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 144


RESULT 11
US-09-746-783-197
```

```
; Sequence 197, Application US/09746783
; Publication No. US20030044935A1
;  GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John M.
;                  LaVallie, Edward R.
;                  Racie, Lisa A.
;                  Treacy, Maurice
;                  Spaulding, Vikki
;                  Agostino, Michael J.
;                  Howes, Steven H.
;                  Fechtel, Kim
;       TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
;                           ENCODING THEM
;       NUMBER OF SEQUENCES: 231
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: MA
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/746,783
;            FILING DATE: 21-Dec-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Milasincic, Debra J.
;            REGISTRATION NUMBER: 46,931
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 227-7400
;            TELEFAX: (617) 742-4214
;   INFORMATION FOR SEQ ID NO: 197:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 167 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 197:
US-09-746-783-197

  Query Match             25.1%;  Score 165.5;  DB 10;  Length 167;
  Best Local Similarity   32.5%;  Pred. No. 4.7e-10;
  Matches   37;  Conservative   22;  Mismatches   50;  Indels    5;  Gaps    3;

Qy           9 TKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFID 68
               :::  | |: | || | : :   ::: | |    |  | ||    : :| ||:| | || ::
Db          54 SRVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNDMFLFKESRITRALVQIVKGLKYMLE 113

Qy          69 FVARETTCSKESNEELTESCE---TKKLGQSLDCNAEVYVVPWEKKI-YPTVNC 118
                  |||  |   :  | : |:      | |:| | :||:|||| :    | : |
Db         114 VEIGRTTCKKNQHLRL-DDCDFQTNHTLKQTLSCYSEVWVVPWLQHFEVPVLRC 166


RESULT 12
US-09-775-932-14
; Sequence 14, Application US/09775932
; Patent No. US20020137671A1
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 121
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-14

  Query Match             21.0%;  Score 138.5;  DB 9;  Length 121;
  Best Local Similarity   31.5%;  Pred. No. 3e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy           8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
               | :   ||  ||:  : |::::     :    |  :|: :||:  :: ||: |:||| |||:
Db           2 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 61

Qy          68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                  | | |       : :|| :|       |   | |: || ||||:
Db          62 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 108


RESULT 13
US-09-775-932-12
; Sequence 12, Application US/09775932
; Patent No. US20020137671A1
; GENERAL INFORMATION:
;  APPLICANT:  University of British Columbia
;  TITLE OF INVENTION:  Production and use of Modified Cystatins
;  FILE REFERENCE:  58069
;  CURRENT APPLICATION NUMBER: US/09/775,932
;  CURRENT FILING DATE:  2001-02-02
;  PRIOR APPLICATION NUMBER:  CA99/00717
;  PRIOR FILING DATE:  1999-08-05
;  PRIOR APPLICATION NUMBER:  60/095,503
;  PRIOR FILING DATE:  1998-08-05
;  NUMBER OF SEQ ID NOS:  32
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-775-932-12

  Query Match             21.0%;  Score 138.5;  DB 9;  Length 128;
  Best Local Similarity   31.5%;  Pred. No. 3.2e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy           8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
               | :   ||  ||:  : |::::     :    |  :|: :||:  :: ||: |:||| |||:
Db           9 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 68

Qy          68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                  | | |       : :|| :|       |   | |: || ||||:
Db          69 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 115


RESULT 14
US-09-940-497-2
; Sequence 2, Application US/09940497
; Patent No. US20020052476A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: Human Cystatin E
;  FILE REFERENCE: PF202P1D2
;  CURRENT APPLICATION NUMBER: US/09/940,497
;  CURRENT FILING DATE:  2001-08-29
;  PRIOR APPLICATION NUMBER: US 09/241,376
;  PRIOR FILING DATE: 1999-02-02
;  PRIOR APPLICATION NUMBER: US 08/744,138
;  PRIOR FILING DATE: 1996-11-05
;  PRIOR APPLICATION NUMBER: US 08/461,030
;  PRIOR FILING DATE: 1995-06-05
```

```
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-940-497-2

  Query Match             21.0%;  Score 138.5;  DB 9;  Length 149;
  Best Local Similarity   31.5%;  Pred. No. 3.9e-07;
  Matches   34;  Conservative   22;  Mismatches   45;  Indels    7;  Gaps    3;

Qy          8 PTKICVGCPRDIPTNSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFI 67
              |  :  ||  ||:  : |::::     :   |  :|: :||:  :: ||: |:||| |||:
Db         30 PQERMVGELRDLSPDDPQVQKAAQAAVASYNMGSNSIYYFRDTHIIKAQSQLVAGIKYFL 89

Qy         68 DFVARETTCSKE----SNEELTESCETKKLGQ--SLDCNAEVYVVPWE 109
                    | | |      : :|| :|       |   | |: || ||||:
Db         90 TMEMGSTDCRKTRVTGDHVDLT-TCPLAAGAQQEKLRCDFEVLVVPWQ 136


RESULT 15
US-08-849-303-16
; Sequence 16, Application US/08849303
; Publication No. US20030221209A1
;  GENERAL INFORMATION:
;    APPLICANT:  Atkinson, Howard J.
;    APPLICANT:  McPherson, Michael J.
;    APPLICANT:  Urwin, Peter E.
;    TITLE OF INVENTION:  MODIFIED PROTEINASE INHIBITORS
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/849,303
;      FILING DATE:  21-MAY-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  1321-1-003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  112 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-849-303-16

  Query Match             20.6%;  Score 136;  DB 8;  Length 112;
  Best Local Similarity   30.6%;  Pred. No. 5.1e-07;
  Matches   34;  Conservative   22;  Mismatches   31;  Indels   24;  Gaps    5;

Qy         22 NSPELEETLTHTITKLNAENNATFYFKIDNVKKARVQVVAGKKYFIDFVARETTCSKESN 81
              |   ::| |:  ::: |  :|  :  ::  | :|| |||:|  ||:|     |||:|  :
Db         12 NEEGVQEALSFAVSEFNKRSNDAYQSRVVRVVRARKQVVSGMNYFLDVELGRTTCTK--S 69

Qy         82 EELTESC--------ETKKLGQSLDCNAEVYVVPWEKKIYPTVN-----CQ 119
              :   :||        : :||      |: :||||||       |:|      ||
Db         70 QANLDSCPFHNQPHLKREKL-----CSFQVYVVPWMN----TINLVKFSCQ 111


Search completed: September 24, 2004, 14:13:03
Job time : 49.296 secs
```